Publication Series of the John von Neumann Institute for Computing (NIC)
NIC Series Volume 39

John von Neumann Institute for Computing (NIC)

NIC Symposium 2008

edited by
Gernot Münster
Dietrich Wolf
Manfred Kremer

Symposium, 20 - 21 February 2008
Forschungszentrum Jülich
Proceedings

organized by
John von Neumann Institute for Computing

NIC Series Volume 39

ISBN: 978-3-9810843-5-1

Die Deutsche Bibliothek – CIP-Cataloguing-in-Publication-Data
A catalogue record for this publication is available from Die Deutsche Bibliothek.

Cover pictures: We thank H. Homann (Bochum) and U. H. E. Hansmann (Jülich) for the pictures used in the cover design.

Publisher:	NIC-Directors
Technical Editor:	Martina Kamps
Distributor:	NIC-Secretariat Research Centre Jülich 52425 Jülich Germany
Internet:	www.fz-juelich.de/nic/
Printer:	Graphische Betriebe, Forschungszentrum Jülich

NIC Series Volume 39
ISBN: 978-3-9810843-5-1

Preface

Gernot Münster
Institut für Theoretische Physik, Universität Münster
Wilhelm-Klemm-Str. 9, 48149 Münster, Germany
E-mail: munsteg@uni-muenster.de

Dietrich Wolf
Fachbereich Physik, Universität Duisburg-Essen, 47048 Duisburg, Germany
E-mail: dietrich.wolf@uni-due.de

Manfred Kremer
John von Neumann-Institut für Computing, 52425 Jülich, Germany
E-mail: m.kremer@fz-juelich.de

Two years have passed since the last NIC Symposium. In the meantime High Performance Computing in Germany has seen developments of great importance for the John von Neumann Institute for Computing (NIC) and its users, and we shall mention some of them below. Computer-based simulations have more and more proved to be highly significant and indispensable in many fields of contemporary research. The NIC symposium and its proceedings provide an overview over this fascinating field.

The research carried out on the computers at NIC ranges from basic to applied sciences. Basic research, with its aim for the acquisition of scientific knowledge, includes among others the physics of elementary particles and astrophysics, both of which are strongly represented at NIC. Monte Carlo calculations with the help of supercomputers lead to a better understanding of the fundamental constituents of matter and their interactions, and of the dynamics of astrophysical systems from the early universe to stars and clusters of galaxies. Closer to applications are projects from materials science and chemistry to earth and environmental sciences and medicine. New types of materials, like nanostructured solids, are investigated and developed by means of numerical simulations. Successful research concerning the various aspects of our environment depends crucially on the availability of the most powerful computers.

The NIC Symposium 2008 brings together scientists from a variety of scientific fields, who have used the facilities of the John von Neumann Institute for Computing (NIC) for their high performance computing applications. Most of the projects come from universities. For the present proceedings of the NIC Symposium a referee panel selected 45 contributions which illustrate the progress in computer based science at NIC. Out of these, 15 projects were selected to be presented in talks at the NIC Symposium.

The prosperity of computer-based simulation sciences does not only depend on the allocation of resources at centres like NIC, but more generally on the presence of a pyramidal structure of computational resources, from local workstations and smaller clusters, high performance computers at universities, up to national and international supercomputer

centres. This pyramidal structure has been specified in a strategy paper by a working group of the BMBF (Bundesministerium für Bildung und Forschung) conducted by Prof. A. Reuter. As a first step towards an organisational implementation of this structure, the three centres at the top of the pyramid – NIC, HLRS Stuttgart and LRZ München – have confederated in the Gauss Centre for Supercomputing (GCS). Its present chairman is Prof. A. Bachem, also chairman of the board of FZ Jülich. With the foundation of the Gauss Centre the discussion about the representation of Germany in the initiative towards the development of a European supercomputer infrastructure has come to a conclusion. The Gauss Centre shall act as the German partner in the framework of the European activities and will represent the German interests. Consequently the Gauss Center participates in the European consortium PRACE.

The "Partnership for Advanced Computing in Europe" (PRACE) has been founded in 2007. Its aim is to prepare an application for the advancement of supercomputer centres in the 7th EU framework programme for research and technology development, and to establish a supercomputer infrastructure on a European level. The intended actions are based on the recommendations of the "HPC for Europe Task Force" (HET). PRACE has five principal partners (Germany, United Kingdom, France, Spain, Netherlands) and further 9 general partners (Finland, Greece, Italy, Norway, Austria, Poland, Portugal, Sweden, Switzerland).

Altogether NIC has further consolidated its role as a leading European supercomputer centre. An important aspect in this respect is the enhancement of supercomputer resources at FZ Jülich, which are available for the NIC community. The new era of capability computing in Jülich, which was started with the IBM Blue Gene/L system in 2006 has been carried forward with the installation of a Blue Gene/P with a peak performance of 230 TFlop/s at the end of 2007. As the successor of the Blue Gene/L it will be available for flagship projects which are particularly suited for such a machine and can use the very large numbers of processors in parallel. The majority of applications at NIC will, however, still be carried out on the JuMP system. At DESY in Zeuthen the parallel computer apeNEXT has been completed with 8 crates of 256 processors each. It is designed in view of simulations of models of elementary particle theory and is used by such projects.

A representative sample of the exciting research, which is done with the help of such machines, is presented at the NIC Symposium 2008. We thank the referees for selecting the contributions for his book. Most of them are also members of the NIC peer review board. We also thank the authors for representing the NIC by their contributions in this book. Special thanks go to the Jülich Supercomputing Centre JSC (the former ZAM) and to DESY for providing the high quality computing environments and user service, which are crucial for the successful realisation of the projects presented in this book and the talks at the symposium. Thanks go also to Ms. Martina Kamps for her effort to compile all the contributions and to create a high quality book from them. Last, but not least we thank Ms. Helga Frank and Ms. Erika Wittig for their help in organizing the 4th NIC-Symposium.

Jülich, February 2008

G. Münster D. Wolf M. Kremer

Contents

Chemistry

Elementary Particle Physics

Materials Science

Recent Developments in Supercomputing

Thomas Lippert

Jülich Supercomputing Centre, Forschungszentrum Jülich
52425 Jülich, Germany
E-mail: th.lippert@fz-juelich.de

Status and recent developments in the field of supercomputing on the European and German level as well as at the Forschungszentrum Jülich are presented. Encouraged by the ESFRI committee, the European PRACE Initiative is going to create a world-leading European tier-0 supercomputer infrastructure. In Germany, the BMBF formed the Gauss Centre for Supercomputing, the largest national association for supercomputing in Europe. The Gauss Centre is the German partner in PRACE.

With its new Blue Gene/P system, the Jülich supercomputing centre has realized its vision of a dual system complex and is heading for Petaflop/s already in 2009. In the framework of the JuRoPA-Project, in cooperation with German and European industrial partners, the JSC will build a next generation general purpose system with very good price-performance ratio and energy efficiency.

1 Introduction

Supercomputers enable scientists and engineers to solve computational problems of unprecedented complexity. Today, computer simulation is established as a third category of gaining scientific insight, in addition to theory and experiment, and has been developed into a key technology for industrial product development and production optimization.

The supercomputers provided by the *Jülich Supercomputing Centre* (JSC), the former Zentralinstitut für Angewandte Mathematik, are utilized by scientists all over Germany and Europe via provision of computer time through the John von Neumann Institute for Computing (NIC). Currently, the Blue Gene/P system at JSC delivers more than 220 Teraflop/s highly scalable supercomputing power. JUGENE is the most powerful machine for free research worldwide and ranks at position 2 in the TOP500 list of November 2007. Supercomputers in this performance class are at the top of an indispensable large-scale scientific infrastructure for national and European research. The JSC strives to guarantee optimal user support, continuously developing the simulation methodology, parallel algorithms and new programming and visualisation techniques as well as carrying out intensive own research in core areas of the computational sciences.

The JSC will intensify its support, research and development activities: support and research are strengthened by establishing several simulation laboratories in the core scientific fields served by the NIC. Each simulation lab is targeted at a specific disciplinary scientific area. Currently the SLs “Plasma Physics”, “Computational Biology” and “Earth System Sciences” are in their first stage of implementation. Other labs like “Nano and Materials Science” will follow soon. While the evaluation and benchmarking of novel compute architectures, in particular in the run-up for a next procurement, have a long tradition at JSC, a novel activity of JSC will be its active participation in the design and building of next generation supercomputer systems (Sec. 4.3).

Similar activities will become an important element of the EU project PRACE, the Partnership for Advanced Computing in Europe (Sec. 3), which is coordinated by Achim Bachem at Forschungszentrum Jülich. PRACE is an initiative of 14 European countries to create a so-called tier-0 supercomputer infrastructure in the Multi-Petaflop/s range that is on one level with the US infrastructure or might even lead the worldwide supercomputing efforts for simulation science and engineering.

In order actively participate in these developments a new association was formed by initiative of the BMBF in Germany (Sec. 2). In June 2006, the Minister for Research, Dr. Schavan, announced the creation of the *Gauss Centre for Supercomputing* (GCS). The GCS-Verein has Forschungszentrum Jülich, Höchstleistungsrechenzentrum Stuttgart and Leibniz-Rechenzentrum Garching as its members. Further members are the chairs of the scientific councils, *e.g.* the NIC for the FZJ. Currently, FZJ is chairing the GCS.

In this note, I am going to review the status and the goals of PRACE and GAUSS. In the framework of PRACE, a technology platform for development of systems in the Multi-Petaflop-range will be created. Currently, consortia like PROSPECT and TALOS are being established jointly by IT-industry and user centres. They are going to support the European technology platform of PRACE. I will describe JSC's approach to Petacomputing, which is guided by our wish to take the users with us on our way to unprecedented scalability. I will finally comment on our efforts with the JuRoPA project (Jülich Research on Petaflop Architectures) to design and to build ourselves a general purpose computer system in the Petaflop-range in cooperation with our industrial partners from Germany and Europe.

2 Gauss Centre for Supercomputing

In order to continue to provide computing resources at the highest performance level for computer-based science and engineering in Germany, in June 2006, the Federal Ministry of Education and Research (BMBF) has initiated the foundation of the Gauss Centre for Supercomputing[1]. The Ministries of Innovation, Science, Research and Technology of the State of North Rhine-Westphalia, the Bavarian State Ministry of Science, Research and the Arts, and the Ministry of Science, Research and the Arts Baden-Württemberg unreservedly support the GCS. Since September 2007, the GCS is a registered association (e.V.).

The three German national supercomputing centres joined up to form the GCS with combined computing power of currently more than 350 teraflop/s, located in Stuttgart, Garching, and Jülich. The GCS is the largest national association for high-performance computing in Europe. It is planned to increase the overall performance of the GCS to a value larger than 1000 Teraflop/s in 2009.

The GCS members follow a common direction in their organization. The procurement of hardware will be closely coordinated, applications for computing time will be scientifically evaluated on a common basis, and software projects will be jointly developed. The GCS currently is preparing documents for a business plan and the future scientific governance of supercomputing in Germany. This task involves also the planning of a closer coordination with the German Tier-1 and Tier-2 centres in an HPC alliance of supercomputer centres.

Another important area of activities of the GCS will be the training of users and the education of young scientists. The work of specialist researchers will be supported and promoted by harmonizing the services and organizing joint schools, workshops, and con-

ferences on simulation techniques. Methodologically oriented user support is also a major concern of the Gauss Centre.

With the establishment of the GCS and its activities, the three GCS sites have already created a high visibility worldwide and have laid the grounds to play a central role in the establishment of a European high-performance supercomputer infrastructure within the PRACE initiative.

3 Partnership for Advanced Computing in Europe

Scientists and engineers in Europe must be provided with access to leadership class supercomputing in order to remain competitive internationally and to maintain or regain leadership in the simulation sciences. In this spirit, the European Strategy Forum for Research Infrastructures (ESFRI) has identified HPC as a strategic priority concerning the creation of new research infrastructures in Europe. Following the recommendations detailed in the ESFRI Roadmap the European Commission issued a call in the 7th Framework Programme for a preparatory phase for up to 35 European research infrastructures[2].

This development goes back to the preparatory studies of HPCEUR and the High Performance Computing in Europe Taskforce (HET)[3] between 2004 and end of 2006. A consortium of 14 European countries signed a Memorandum of Understanding on April 17, 2007, in Berlin, in the presence of minister Schavan. The initiative is named PRACE, *i.e.* Partnership or Advanced Computing in Europe. Prior to this event, the founding members (HET) have submitted a project proposal, the PRACE project, to meet the above mentioned call. The project was fully granted in summer 2007 and will have its kick-off meeting end of January 2008.

PRACE aims at preparing a persistent pan-European HPC supercomputer service in form of a single legal entity, consisting of three to five tier-0 centres, similar to the US HPC-infrastructure. PRACE will be the tier-0 level of the European HPC ecosystem. The hosting centres of the planned tier-0 systems will provide the expertise, competency, and the required infrastructure for comprehensive services to meet the challenging demands of excellent users from academia and industry.

PRACE will prepare for the implementation of the infrastructure in 2009/2010. Within the project the definition and the set-up of a legal and organizational structure involving HPC centres, national funding agencies, and scientific user communities are carried out. Tasks involve plans to ensure adequate funding for the continued operation and periodic renewal of leadership systems, co-ordinated procurements, efficient use and fair access for European researchers.

PRACE will prepare the deployment of Petaflop/s systems in 2009/2010. This includes the procurement of prototype systems for the evaluation of software for managing the distributed infrastructure, the selection, benchmarking, and scaling of libraries and codes from scientific user communities, the definition of technical requirements and procurement procedures.

The PRACE Petaflop-infrastructure will require an initial investment of up to 500 Mio Euro followed by annual funds of 100 Mio Euro for upgrades and renewal in later years.

3.1 Technical Development in Europe: Consortia of Industry and User Centres

In addition to the precommercial procurement activities for the PRACE Petaflop-systems, a collaboration with the European IT-industry will be started by PRACE in order to influence the development of new technologies and components for architectures that are promising for Petaflop/s systems to be procured after 2010.

In order to achieve these goals, already early in 2007 two consortia were formed in Europe. One consortium is PROSPECT (Promotion of Supercomputers and Petacomputing Technologies), with the partners BSC, JSC, LRZ, DWD, IBM, Quadrics, Intel, ParTec, University of Heidelberg and University of Regensburg. A second one is TALOS centered on BULL and CEA/France, with the associated partners Intel and Quadrics. These consortia, and maybe more to come, will seek cooperation with the workgroup WP8 within the PRACE project in order to create a long-term technology platform for development of systems in the realm of Multi-Petaflop/s.

PROSPECT understands itself as an interest group to foster the development of supercomputing technology in Europe. Partners within PROSPECT have started to form project groups that focus on hardware and software aspects. So far three projects are being initiated within the PROSPECT framework. One of these project is JuRoPA (Sec. 4.3), a second one is called QPACE (QCD Petacomputing Engine). Within QPACE, IBM-Böblingen, the universities of Regensburg and Wuppertal, DESY-Zeuthen and the JSC are going to build a new 3-dimensional high-speed communication network around a grid of next-generation cell processors.

A third project within PROSPECT will involve most of the PRACE partners tackling the highly relevant question of energy aware and energy efficient supercomputing. Energy issues become ever more important[4] since the need for higher performance requires to boost parallelism after frequency growth came to an end. Therefore, the individual processors necessarily must become less instead of more power consuming. The Blue Gene approach as well as the cell processor are state-of-the-art technologies concerning power efficiency. While with the new multi-core processors of Intel and AMD an important step is taken in the right direction, energy consumption still grows linearly with the number of processors, which imposes a severe limit to scalability of future supercomputer hardware. As a consequence, we have to make use of all possible ideas to reduce the waste of power.

4 Jülich Supercomputing Centre

The JSC works at the forefront of the technologically possible in HPC. Following a dual philosophy the JSC strives to provide Petaflop-class systems with highest scalability like the new Blue Gene /P together with highly flexible general purpose systems. In 2008 the JUMP general purpose system will be replaced by a very large compute cluster increasing the power of JUMP by more than a factor of 25. This cluster is currently designed and will be built together with partners from industry. The operating system software to be utilized is the German system ParaStation, developed jointly by ParTec (Munich) and JSC. Together with an online storage of 1 PetaBytes, connected to an automatized tape archive of about 10 Petabytes the JSC operates the most flexible computer complex currently available.

4.1 Dual System Complex

Already quite early in the operation cycle of JUMP, the Parateam at JSC, responsible for the scaling of user applications, realized that for the given number of users and allocation of computing time a further growth of capability computing on JUMP would have decreased the overall efficiency. This is because many users worked at their sweet spot and would not have profited from policies preferring jobs with very large processor numbers.

As a solution, the concept of a "Dual Supercomputer Complex" was conceived and described as part of a white paper of the Helmholtz Association with title "Ausbau des Supercomputing in der Helmholtz-Gemeinschaft und Positionierung im europäischen Forschungsraum".

By installing the first Blue Gene/L system (Sec. 4.2) and boosting it up to 46 Teraflop/s, the JSC started to realize the dual concept. Blue Gene/L is only allowed to be used by highly scaling applications. Before being admitted to the machine, the scalability of a code must be proven. In this way, a decisive step towards capability computing was taken, as is shown by the performance parameters of the Blue Gene/L system. Following this philosophy, the number of accepted projects on JUBL for the granting period 2006/2007 was limited to about 25. A substantial increase of the scalability up to 16,384 processors could be reached for several applications by intensive user support during two workshops, the "Big Blue Gene Week" and the "Blue Gene Scaling Workshop" which both lasted for one week[5].

The dual system complex is supported by a common storage infrastructure which was also expanded in 2007. Key part of this infrastructure is the new Jülich storage cluster (JUST) which was installed in the third quarter of 2007, increasing the online disk capacity by a factor of ten to around one PetaByte. The maximum I/O bandwidth of 20 GB/s is achieved by 29 storage controllers together with 32 IBM Power 5 servers. JUST is connected to the supercomputers via a new switch technology based on 10-Gigabit Ethernet. The system hosts the fileserver function for GPFS (General Parallel File System) and provides service to the clients in Jülich and to the clients within the international DEISA infrastructure as well. In future technological developments, a LUSTRE parallel file system will join the GPFS.

4.2 Towards Petacomputing

In 2004/2005, the IBM Blue Gene technology became available. The JSC has evaluated the potential of this architecture very early as a future leadership class system for capability computing applications. A key feature of this architecture is its scalability towards Petaflop-computing based on lowest power consumption, smallest footprint and an acceptable price-performance ratio.

In early summer 2005 Jülich started testing a single Blue Gene/L rack with 2048 processors. It soon became obvious that many more applications than expected could be ported to efficiently run on the Blue Gene architecture. Due to the fact that the system is well balanced in terms of processor speed, memory latency and network performance, many applications scale up to large numbers of processors. In January 2006 the system was expanded to 8 racks with altogether 16384 processors.

The 8-rack system has successfully been in operation for almost two years now. About 4 racks or about 25 Teraflop/s are available for the users in the NIC. About 25 research

projects, which were carefully selected with respect to their scientific quality, run their applications on the system using job sizes between 1024 and 16384 processors. During the Blue Gene Scaling Workshop at FZJ, experts from Argonne National Laboratory, IBM and Jülich helped to further optimize some important applications. As already stated, it could be shown that all applications considered succeeded in efficiently using all 16384 processors of the machine[5].

Computational scientists from many research areas took the chance to apply for significant shares of Blue Gene/L computer time to tackle unresolved questions which were out of reach before. Because of the large user demand and in line with its strategy to strengthen leadership-class computing, FZJ decided to order a powerful next-generation Blue Gene system. In October 2007 a 16-rack Blue Gene/P system with 65536 processors was installed mainly financed by the Helmholtz Association (BMBF) and the State of North Rhine Westphalia. With its huge peak performance of 222.8 TFlop/s, Jülich's Blue Gene/P dubbed "JUGENE" is currently the most powerful supercomputer in Europe and number 2 worldwide.

The important differences between Blue Gene/P and Blue Gene/L mainly concern the processor and the networks (see Table 1) while the principal architecture of Blue Gene/L was kept unchanged. Key features of Blue Gene/P are:

- 4 PowerPC 450 processors are combined in a fully 4-way SMP (node) chip which allows a hybrid programming model with MPI and OpenMP (up to 4 threads per node).
- The network interface is fully DMA (Direct Memory Access) capable which increases the performance while reducing the processor load during message handling.
- The available memory per processor has been doubled.
- The external I/O network has been upgraded from 1 to 10 Gigabit Ethernet.

All these improvements are well reflected by the application performance. A code from theoretical elementary particle physics, (dynamical overlap fermions), on Blue Gene/P runs at 38% of the peak performance compared to 26.3% on Blue Gene/L. Furthermore, the increased memory of 2 GB per node will allow new applications to be executed on Blue Gene/P. JUGENE has become part of the dual supercomputer complex of JSC. It will be complemented soon by the successor of the JUMP system, a general purpose system based on cluster technology (see Sec.4.3).

With the upgrade of its supercomputer infrastructure FZJ has taken the next step towards Petascale Computing and has strengthened Germany's position to host one of the future European supercomputer centres.

4.3 Building a Supercomputer

In the future, the JSC will play an active role in designing and building its supercomputers, in particular as far as general purpose systems are concerned. Using component-oriented cluster technology JSC will be able to directly address its users' requirements in the design and build-up process and to go for the best possible price-performance ratio and lowest energy consumption at the earliest possible implementation date.

	Blue Gene/L	**Blue Gene/P**
Node Properties		
Processor	PowerPC 440	PowerPC 450
Processors per node (chip)	2	4
Processor clock speed	700 MHz	850 MHz
Coherency	Software managed	SMP
L1 cache (private)	32 KB per core	32 KB per core
L2 cache (private)	7 stream prefetching	7 stream prefetching
	2 line buffers/stream	2 line buffers/stream
L3 cache (shared)	4 MB	8 MB
Physical memory per node	512 MB	2 GB
Main memory bandwidth	5.6 GB/s	13.6 GB/s
Peak performance	5.6 GFlop/s	13.6 GFlop/s
Torus network		
Bandwidth	2.1 GB/s	5.1 GB/s
Hardware latency (nearest neighbour)	200 ns (32B packet)	160 ns (32B packet)
	1.6 μs (256 B packet)	1.3 μs (256 B packet)
Global collective network		
Bandwidth	700 MB/s	1700 MB/s
Hardware latency (round trip worst case)	5.0 μs	3.0 μs

Table 1. Blue Gene /P vs. Blue Gene /L.

At this stage, JSC is in the process to design the successor of JUMP. It is planned to realize a highly flexible general purpose cluster system comprising 2048 compute nodes to be installed end of 2008. Each node will host 8 cores of the next generation Intel HPC processor Nehalem. The high-speed interconnect will be based on QSnet^{III} of the British-Italian IT-company Quadrics. As cluster operating system the German software ParaStation will be used which is proven to scale beyond 1000 nodes.

To this end, a research project group was formed named JuRoPA (Jülich Research on Petaflop Architectures). The core partners of JuRoPA are the JSC, ParTec (Munich)[6], Intel (Brühl)[7], and Quadrics (Bristol)[8]. At a later stage, a provider for the hardware which is not yet selected will join the group.

The ParaStation cluster operating system is being developed in cooperation between ParTec and JSC since several years. Together with the European interconnect technology Quadrics which will again become the state-of-the-art technology in the field of cluster computing with its next generation QSnet^{III} the JSC will demonstrate that Europe is back in supercomputing technology and can build leadership-class supercomputers at lowest costs and high energy efficiency.

5 Concluding Remarks

In the last two years, coordinated efforts in Europe and Germany have led to a variety of new developments in the ever growing field of simulation science and engineering, and in particular in the field of supercomputing. With PRACE, the Partnership for Advanced

Computing, Europe is on the right track to create a Petacomputing infrastructure which can be competitive with US installations. The GCS, the Gauss Centre for Supercomputing, brings the German National Supercomputer Centres in Garching, Jülich and Stuttgart together, forming the largest supercomputing complex in Europe and striving to host the first European supercomputer centre with Petaflop capability.

The JSC, the Jülich Supercomputing Centre, has realized its vision of a dual supercomputing complex both in terms of hardware and computer time provision, which is coordinated by the John von Neumann Institute for Computing. With 223 Teraflop/s, the most powerful supercomputer for free research worldwide is accessible by the NIC community. Together with the successor of the general purpose system JUMP, a cluster being designed and built by JSC, ParTec, Intel and Quadrics, and planned to be in operation end of 2008, German and European scientists can work at the forefront of simulation science.

Acknowledgments

Th.Lippert thanks N. Attig, A. Bachem, J. Docter, N. Eicker, Th. Eickermann, R. Esser, W. Grich, S. Hfler-Thierfeldt, B. Orth, S. Krieg, M. Stephan, and K. Wolkersdorfer for kind contributions to this note.

References

1. http://www.gauss-centre.eu.
2. http://cordis.europa.eu/esfri/.
3. http://www.hpcineuropataskforce.eu.
4. http://www.green500.org.
5. Wolfgang Frings, Marc-André Hermanns, Bernd Mohr, and Boris Orth (Editors), *Report on the Jülich Blue Gene/L Scaling Workshop 2006*, http://www.fz-juelich.de/jsc/docs/printable/ib/ib-07/ib-2007-02.pdf, FZJ-ZAM-IB-2007-02.
6. http://www.cluster-competence-center.com.
7. http://www.intel.com/jobs/germany/.
8. http://www.quadrics.com/quadrics/QuadricsHome.nsf/DisplayPages/Homepage.

Protein Simulations on Massively Parallel Computers

Jan H. Meinke[1], Sandipan Mohanty[1], Walter Nadler[1,2], Thomas Neuhaus[1], Olav Zimmermann[1], and Ulrich H.E. Hansmann[1,2]

[1] John v. Neumann Institute for Computing (NIC), Research Centre Jülich, 52425 Jülich, Germany
E-mail: u.hansmann@fz-juelich.de

[2] Dept.of Physics, Michigan Technological University, Houghton, MI 49931, USA

We summarize shortly the research done in the "Computational Biology and Biophysics" group at NIC. This research group was founded in July 2005 and spearheads the use of modern supercomputers for research into the Biophysics and Biochemistry of biological macromolecules.

1 Introduction

The need for high performance computing in Biochemistry and Systems Biology is well-recognized. For most sequences in the now deciphered genomes, the structures and functions of the corresponding proteins are not known. Even less is known about their interaction and regulation. Reliable tools that allow one to study these phenomena in computer experiments would open the way for understanding the molecular foundations of the workings of whole cells. This is because proteins are the workhorses in a cell, transporting molecules, catalyzing biochemical reactions, or fighting infections. Hence, improved computational tools could lead to a deeper understanding of various diseases that are caused by the mis-folding of proteins, and enable the design of novel drugs with customized properties.

Computer simulations of even small proteins have remained a challenging computational task. This is because the complex form of the forces within and between molecules leads to a rough energy landscape with a huge number of local minima acting as traps. As a rule, computational cost to accurately calculate physical quantities in simple room-temperature Molecular Dynamics or Monte Carlo simulations increases exponentially with the number of residues. Overcoming these obstacles requires both new sampling techniques and the use of modern massive parallel computers (such as the new BlueGene/P computer JUGENE in Jülich) that soon will approach the Petaflop range.

Research in the group is concerned with the development of algorithms, and their implementation into software optimized for modern supercomputer architectures, for overcoming these numerical difficulties in protein studies. Our aim is to develop protocols that allow atomistic simulations of stable domains in proteins (usually of order 50-200 residues). Successful method and software development is at its best when it is guided by the needs of specific applications. For this reason, we have studied extensively the folding mechanism in a number of small proteins. Finally, we are interested in the combination of our techniques with knowledge-based approaches. Such combinations are the backbone of the structure prediction algorithms that we have tested successfully at the CASP7 (Critical Assessment of protein Structure Prediction) competition in summer 2006. Our newly developed techniques are included in the recently published update of SMMP[1,2], the freely available program package co-developed by this group.

2 Algorithms for Protein Simulations

The key-idea behind our novel techniques is to replace canonical simulations, where crossing of an energy barrier of height ΔE is suppressed by a factor $\propto \exp(-\Delta E/k_B T)$ (k_B is the Boltzmann constant and T the temperature of the system), with schemes that both ensure sampling of low-energy configurations *and* avoid trapping in local minima. For instance, in multicanonical sampling[3] the weight $w(E)$ in a Monte Carlo or molecular dynamics simulation is set so that the distribution of energies $P(E)$ is given by:

$$P(E) \propto n(E)w(E) = \text{const}, \tag{1}$$

where $n(E)$ is the spectral density. In this way, a free random walk in the energy space is performed that allows the simulation to escape from any local minimum. The thermodynamic average of a physical quantity A can now be calculated by re-weighting:[4]

$$< \mathcal{A} >_T \;\; = \;\; \frac{\int dx\; \mathcal{A}(x)\; w^{-1}(E(x))\; e^{-E(x)/k_B T}}{\int dx\; w^{-1}(E(x))\; e^{-E(x)/k_B T}} \; . \tag{2}$$

Here, x stands for configurations. Note that the weights $w(E)$ are not *a priori* known, and estimators have to be determined by an iterative procedure described in Refs.[3,5].

In parallel tempering (also known as replica exchange method)[6], first introduced to protein science in Ref.[7], standard Monte Carlo or molecular dynamics moves are performed in parallel at different values of a control parameter, most often the temperature. At certain times the current conformations of replicas at neighbouring temperatures T_i and $T_{j=i+1}$ are exchanged with probability

$$w(\mathbf{C}^{old} \rightarrow \mathbf{C}^{new}) = \min(1, \exp(-\beta_i E(C_j) - \beta_j E(C_i) + \beta_i E(C_i) + \beta_j E(C_j))) \; , \tag{3}$$

with $\beta = 1/k_b T$. For a given replica the swap moves induce a random walk from low temperatures, where barriers lead to long relaxation times, to high temperatures, where equilibration is rapid, *and back*. This results in a faster convergence at low temperatures.

However, these sophisticated novel techniques are hampered still by slow relaxation due to barriers and bottlenecks. Their successful application to protein simulations requires that the weight functions or temperature distributions are chosen optimally. In order to unify the analysis of these techniques, we derived equivalent one-dimensional stochastic processes from the underlying Master equations and analyzed the conditions under which these representations are valid descriptions of the random walk in order parameter or control parameter space[8] . They allow a unified discussion of the stationary distribution on each space, as well as of the stationary flow across it. We demonstrated that optimizing the flow is equivalent to minimizing the first passage time for crossing the space, and discussed the consequences of our results for optimizing simulations, particularly under conditions of broken ergodicity. Based on these results we were able to determine an analytical expression for the optimal number of replicas in PT simulation[9]. In addition, we have investigated the theoretical basis for combining molecular dynamics simulation with parallel tempering and developed optimized replica exchange move sets for that[10].

Another example of our algorithmic research is concerned with temperature driven first order phase transition where traditional parallel tempering implementations fail. We have developed an efficient new parallel tempering algorithm[11], that eliminates the supercritical slowing down associated with the nucleation barrier at the transition and which is capable

of determining the density of states function. Our algorithm is much simpler than multi-canonical ensemble simulations, which on the input side need an a-priori unknown weight function - or - Wang Landau simulations, which require a tedious parameter fine tuning. In addition, multiple Gaussian modified ensemble simulations[11] are perfectly suited for parallel computer architectures. Future applications of the method will include studies of condensates with ensembles of chain molecules.

3 Protein Simulation Programs for Parallel Computers

The above algorithms are implemented in SMMP[1], our program package for simulation of protein. The code is free and open source. The latest version[2] features a Python interface and allows simulations of more than one protein as necessary for studies of aggregation or ligand binding. The implementation of additional force fields are other newly added features. SMMP is available from either the program library of *Computer Physics Communications* or from *www.phy.mtu.edu/biophys/smmp.htm*

Emphasis was put on the parallelization of SMMP: we now regularly run our simulation on 4096 processors. In SMMP, every atom is associated with a dihedral angle. We used this relation to distribute the interactions as evenly across processors as possible without regard to spatial proximity. We ran our benchmark on 4 different platforms: JUMP, an IBM p690 cluster with 32 Power4+ processors at 1.7 GHz and 112 GB of shared memory per frame and a total of 1312 processors; JUBL, an IBM BlueGene/L with 8 racks and a total of 16384 Power4 processor at 700 MHz; JULI a PC cluster using dual-core PowerPC 970MP processors at 2.5 GHz with an InfiniPath network and NICOLE, an Opteron based PC cluster with a clock speed of 2.4 GHz using Infiniband networking. Except for the setup of the communicators used for the energy calculation on BG/L, we used the same source code for all measurements. We performed 50 sweeps of a Monte Carlo simulation of the designed protein Top7 (92 residues, 1477 atoms) starting from a stretched chain. Data was written to disk every 10 sweeps. On JUBL, we used multiple replicas in parallel with the indicated number of processors per replica to fill a half plane (512 processors). Figure 1 shows walltime and scaling for the various machines. The execution time on a single processor ranges from about 18 min on JULI to about 2 h on JUBL. The lowest execution time ranges from 81 s on JUMP to 269 s on JUBL with 64 processors per replica. The maximum speedup is 25 on JUBL with 64 processors.

For JUMP, JULI, and NICOLE we used MPI's default processor assignment. On JUBL, however, this approach leads to a sub-optimal distribution of the processors. BG/L has a cubic geometry. By default, the rank of a processors increases first along x, then y, and finally z. This leads to a planar distribution of processors. Instead of the default, one should make communicators as cubic as possible unless the problem geometry suggests a different approach. The low cost per processor makes BlueGene/L an attractive platform for protein simulations. Using a cubic arrangement of 64 processors, we achieve a speedup of up to $25\times$ on BG/L. With the large number of processors available on JUBL, we can run simulations with 64 replicas at a quarter of the cost and at the same speed as on JUMP.

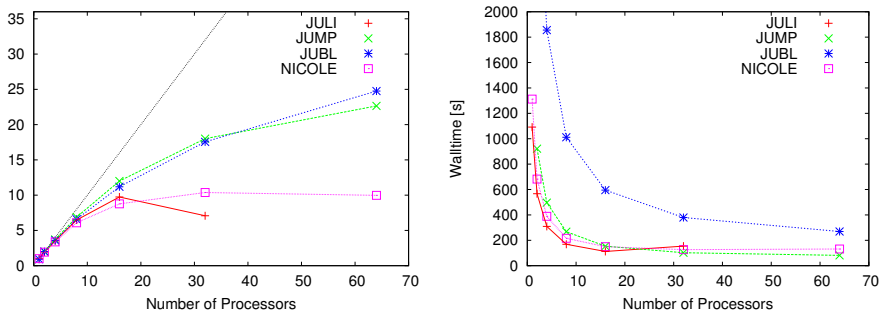


Figure 1. Strong scaling behaviour. This figure shows the walltime and corresponding parallel scaling vs. number of processors on JULI, JUMP, and JUBL. On JUBL, we used multiple replicas in parallel with the indicated number of processors per replica to fill a half plane.

4 Side Chain Ordering in Polymers and Proteins

Work of the group is characterized by the combination of algorithm development with application of these novel techniques to research the physics of proteins and their interaction. One area of interest is the distinct transitions in the folding process, their thermal order and relation. An important example is the role of side-chain ordering.

In recent studies on homopolymers[12, 13], we found for certain amino acids a de-coupling of backbone and side-chain ordering. The characteristics of the side chain ordering process did not depend on the details of the environment, i.e. whether the molecules were in gas phase or solvent, but solely on the side groups. It exhibited a phase transition-like character, marked by an accompanying peak in the specific heat. In a related investigation, we were able to establish the role of charged end groups in stabilizing and de-stabilizing secondary structures in gas phase[14]. These results are important for comparison with molecular beam experiments. Later, we have extended those investigations to proteins, starting with the villin headpiece subdomain HP-36[15]. This molecule has raised considerable interest in computational biology as it is one of the smallest proteins (596 atoms) with well-defined secondary and tertiary structure but at the same time still accessible to simulations. Our results indicate a thermal hierarchy of ordering events with side-chain ordering appearing at temperatures below the helix-coil transition, i.e. secondary structure formation, but above the final folding transition to the native state. We believe that the observed thermal hierarchy of folding reflects an underlying temporal sequence of these ordering processes in actual protein folding dynamics. We conjecture that side-chain ordering facilitates the search for the correct backbone topology. In contrast to homopolypeptides we do not find collective effects leading to a separate transition. The heterogeneity of the sequence seems to destroy the phase transition-like character of side-chain ordering.

5 Folding Mechanisms in Small Proteins

The above investigations were concerned with helical polypeptides. In an α-helix, hydrogen bonds are formed between residues i and $i + 4$ along the sequence. Because of

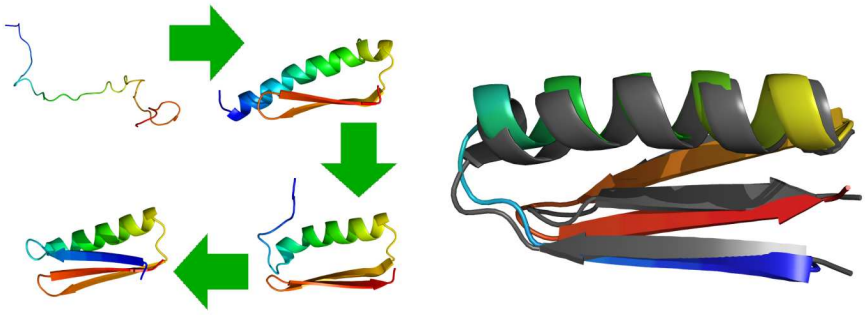


Figure 2. Left: The non-trivial unexpected pathway for the folding of the molecule CFr revealed in the simulations. Right: The free-energy mimimum structure seen in all-atom simulations (coloured) superimposed on the experimentally measured structure (gray).

this short sequence separation, these residues are constrained so that the helical hydrogen bonds form along with chain compactification. Small helical proteins indeed show simple funnel like folding free-energy landscapes[16]. The helix hydrogen bonds form in no particular order, although the two ends of a helix show greater tendency to dissolve and reform. β-hairpins are also "local" structures in that the hydrogen bonded residues are close in sequence. For the 3-stranded β-sheet beta3s, we find the folding of the two β-hairpins more cooperative than the folding of helices. Folding proceeded in a zipper mechanism from the turns towards the ends of the hairpins[16]. Once formed, the β-sheets show a greater resilience towards unfolding.

The simplest example of a protein with both helix and β-sheet elements is a structure with a helix and a β-hairpin. We have examined two such systems and found rather distinct mechanisms. The 23 residue BBA5 molecule has a small β-hairpin where the turn region is stabilized by a synthetic amino acid D-proline. This hairpin and the helix of BBA5 form on their own, and only later make hydrophobic contacts[16]. On the other hand, in a simulation of the protein FSDEY with a similar helix hairpin structure, we found that the hydrophobic residues of the helix line up on one side, to provide a template around which the hairpin forms[17]. The hairpin of FSDEY never forms independently of the helix in our simulations, nor did we find any clear evidence of a zipper like mechanism.

The formation of structures with non-local β-sheet contacts are highly non-trivial. We have done extensive folding simulations of one such molecule, CFr (PDB id: 2GJH), the C-terminal fragment of the designed 93 residue protein Top7 (PDB id: 1QYS). Along the sequence from N- to C- terminal, the secondary structure profile of the molecule CFr is : strand – helix – strand – strand (see Fig.2). The two strands at the C-terminal make a β-hairpin. The strands at the N- and C-termini are also adjacent in the 3 stranded β-sheet. None of the simple folding mechanisms discussed above could give rise to this arrangement. Instead, our simulations revealed an unanticipated mechanism (see Fig. 2) for the formation of this structure[18]. The N-terminal β-strand first folded into a non-native extension of the native helix. The β-hairpin at the C-terminus forms independently. When the helix and the C-terminal hairpin make the correct tertiary contacts, the non-native part of the helix unfolds to release the N-terminal residues. These subsequently make β-sheet

contacts with the hairpin and complete the native structure. By "caching" the N-terminal β-strand as a non-native extension of a helix, the molecule protects them from premature contacts with other regions with strong β-strand propensities, which would lead to misfolding, or very slow folding. The caching of the N-terminal strand, accelerates folding of CFr by avoiding many misfolded states. We speculate that this mechanism is employed in molecules where adjacent strands in a β-sheet have large sequence separation. The same mechanism could protect a nascent N-terminal β-strand which is synthesized early, from intermolecular interactions leading to aggregation, until the rest of the molecule is synthesized and properly folded. Simulations of FSDEY were done with SMMP[2], while we used PROFASI[19] for the other proteins mentioned in this section.

6 Thick Polymers

The results described above raise the question to what extent the properties of proteins can also be explained as polymer properties (i.e. are not sequence specific). To investigate this, we simulated "thick" polymers in a simple $3d$ homo-polymer chain model with thickness regularized by the global radius of curvature, and studied the interplay of "hard" geometric constraints and attractive Lennard Jones interactions. In an earlier work the long chain limit of thick polymers was studied[20] and was found to be consistent with field theoretic expectations. As a function of the global radius of curvature we now find a rich "landscape" of low temperature i.e., ground state configurations, which we classify. These include simple hypercubic crystals, ideal helices as introduced by Maritan[21], short sheet like structures, twisted circles and helical superstructures. We also observe a phase region where open ended polymers close into ring polymers with zero knot number and with short range interactions as being characteristic for a liquid. While these studies are of principal nature, some similarities with real proteins are seen. In particular we observe secondary structure formation as well as mixed phases of few conformational simplices for long chains. However, a direct mapping onto protein configurations is problematic and there is clear need for the inclusion of additional either protein specific - or more fundamental interactions into the simplified model. A possible candidate is a thick polymer model with attractive Lennard Jones interactions as well as dipole-dipole interactions on the chain.

7 Constraints for Structure Prediction

Our algorithms are limited to simulations of small proteins of order ≈ 50 residues. As the average size of proteins is around 250 residues, it is necessary to constrain the conformational search space for the purpose of structure prediction. Such constraints are generally obtained from known structures of similar sequences. As long range distance constraints (e.g. from fold recognition or correlated mutation analysis) lead to frustrated conformations, local constraints e.g. dihedral constraints from secondary structure predictions are preferred. State of the art algorithms like PSIPRED[22] achieve highly accurate secondary structure predictions, but do not allow mapping of these classes to dihedral angles. To overcome these problems we have developed DHPRED to predict for each residue its dihedral angle region[23]. The algorithm has a three layer structure and is based on classification by Support-Vector-Machines (SVM)[24], a supervised machine learning algorithm. The performance of DHPRED is comparable to PSIPRED but provides a direct mapping to dihedral

angles. In addition the dihedral angle regions for many coil residues can be identified thus providing local dihedral constraints for the entire protein chain.

Patterns of dihedral angles i.e. correlations between the dihedral state of neighbouring residues can also be observed in coil regions. Among these patterns β-turns[25] are the most abundant class. They are frequently encountered at topologically interesting positions where the chain changes its direction and as such are important prediction targets. We have developed a classifier to resolve the different classes of β-turns as they indicate different local topologies[26]. While certain classes can be easily distinguished, several other classes (I, IV, VIII) show considerable overlap. As these classes also share certain geometrical properties it is an open question whether this is due to an insufficient number of samples available or to the fact that the class boundaries are artificial.

The rapid growth of the protein databank PDB makes it increasingly likely that a structurally related protein to a sequence has already been entered even if the sequence similarity is too low to allow for direct identification. For these cases all successful structure prediction strategies use fold recognition, where structural features are derived from the target sequence and compared to a structure database. In the simplest case secondary structure alignment is used. As ordinary secondary structure has only three different elements there is a high likelihood that two proteins with the same secondary structure have vastly different topology. A considerable reduction of the possible topologies can be achieved by predicting which β-strands are part of an antiparallel β-sheet and which form a parallel one. Our program BETTY[27] implements an SVM-approach similar to the one described for DHPRED but without the iterative second layer. It classifies 88% of all β-residues correctly. Combined with PSIPRED 79.3% can be correctly classified into parallel-β, antiparallel-β, α-helix, and coil. We are planning to use this enhanced secondary structure alphabet as part of a fold recognition algorithm.

In order to test our techniques for structure prediction we have participated in the CASP 2006 competition where we have submitted structures for all but three of the 100 valid targets. As expected our approach is not competitive in cases where homology to resolved proteins allows application of knowledge-based techniques. However, we are within the top five groups in the free modelling section when ranked based on visual inspection (see `http://predictioncenter.org/casp7/meeting/presentations/Presentations_assessors/CASP7_FM_Clarke.pdf`). We are currently evaluating our protocols in order to determine how we can optimize the use of our techniques for structure prediction, and plan to set up a server that automates the predictions for the 2008 round of the CASP competition.

8 Closing Remarks

We have outlined research done by the "Computational Biology and Biophysics" group since its inception in July 2005. Using high performance computing and developing novel algorithms (the "generalized-ensemble" approach) the group has made substantial progress on the way toward the goal of structure prediction of stable domains in proteins (usually of order 50-200 residues). In future, the group will extend these lines of research to larger and medically relevant proteins.

Acknowledgments

Part of the presented work was supported also through research grants of the National Science Foundation (CHE-0313618) and the National Institutes of Health (GM62838).

References

1. F. Eisenmenger, U.H.E. Hansmann, Sh. Hayryan, C.-K. Hu, *Comp. Phys. Comm.* **138** (2001) 192; *Comp. Phys. Comm.* **174** (2006) 422.
2. J. H. Meinke, S. Mohanty, F. Eisenmenger and U. H. E. Hansmann, *Comp. Phys. Comm.*, (2007), doi: 10.1016/j.cpc.2007.11.004.
3. B. Berg and T. Neuhaus, *Phys. Lett.* **B267** (1991) 249; B. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68** (1992) 9.
4. A.M. Ferrenberg and R.H. Swendsen, *Phys. Rev. Lett.* **61** 2635 (1988); **63** (1989) 1658(E), and references given in the erratum.
5. U.H.E. Hansmann and Y. Okamoto, *Physica A* **212** (1994) 415.
6. K. Hukushima and K. Nemoto, *J. Phys. Soc. (Japan)*, **65** (1996) 1604; G.J. Geyer, *Stat. Sci.* **7** (1992) 437.
7. U.H.E. Hansmann, *Chem. Phys. Lett.* **281** (1997) 140.
8. W. Nadler and U.H.E. Hansmann, Phys. Rev. E, **75** (2007) 026109.
9. W. Nadler and U.H.E. Hansmann, Phys. Rev. E, **76**, (2007) 065701(R).
10. W. Nadler and U.H.E. Hansmann, Phys.Rev. E **76** (2007) 057102.
11. T. Neuhaus, M. Magiera and U. Hansmann, Phys. Rev. E **76**, (2007) 045701(R).
12. Y. Wei, W. Nadler and U.H.E. Hansmann, J. Chem. Phys. **125** (2006) 164902.
13. Y. Wei, W. Nadler and U.H.E. Hansmann, J. Phys. Chem. B **111** (2007) 4244.
14. Y. Wei, W. Nadler and U.H.E. Hansmann, J. Chem. Phys., **126** (2007) 204307.
15. Y.Wei, W. Nadler and U.H.E. Hansmann, Backbone and Sidechain Ordering in a small Protein, J. Chem. Phys., in press.
16. S. Mohanty and U.H.E. Hansmann, *Biophysical Journal* **82** (2006) 3573.
17. S. Mohanty and U.H.E. Hansmann, *J. Chem. Phys.* **127** (2007) 035102.
18. S. Mohanty, J. Meinke, O. Zimmermann and U.H.E. Hansmann, Caching of Chameleon Segments: a New Folding Mechanism, submitted for publication.
19. A. Irbäck and S. Mohanty *J. Comp. Chem.* **27**, (2006) 1548.
20. T. Neuhaus, O. Zimmermann and U. Hansmann, Phys. Rev. E **75**, (2007) 051803.
21. A. Maritan, C. Micheletti, A. Trouato and J. Banavar, *Nature* **406**, (2000) 287.
22. D.T. Jones, J. Mol. Biol. **292** (1999) 195.
23. O. Zimmermann and U.H.E. Hansmann, Bioinformatics **22** (2006) 3009.
24. B. Schökopf and A.J. Smola, Regularization, Optimization, and Beyond, MIT Press, Cambridge, MA, 2002.
25. P.N. Lewis, F.A. Monany, and H.A. Scheraga, Biochem. Biophys. Acta **303** (1973) 211.
26. O. Zimmermann and U.H.E. Hansmann, Dihedral Angle Patterns in Coil Regions of Protein Structures, in: Proceedings of CBSB07, U.H.E. Hansmann, J.H. Meinke, S. Mohanty, O. Zimmermann (Eds.), John v. Neumann Inst. for Computing, NIC-Series, vol. 36, Jülich, Germany, 2007, pp. 301–304.
27. O. Zimmermann, L. Wang and U.H.E. Hansmann, In Silico Biol. **7** (2007) 0037.

Phase Transitions and the Mass of the Visible Universe

Zoltan Fodor

John von Neumann Institute for Computing (NIC), Research Centre Jülich, 52425 Jülich, Germany
DESY, Zeuthen, Platanenallee 6, 15738 Zeuthen, Germany
E-mail: fodor@bodri.elte.hu

Recent results on QCD thermodynamics are presented. The nature of the T>0 transition is determined, which turns out to be an analytic cross-over. The absolute scale for this transition is calculated. In order to approach the continuum limit four different sets of lattice spacings were used with temporal extensions N_t=4, 6, 8 and 10 (they correspond to lattice spacings $a\sim$0.3, 0.2, 0.15 and 0.12 fm).

1 Introduction

One of the most fundamental fields of all science is the physics of elementary particles. We are looking for the smallest building blocks of nature. We would like to understand their interactions. Since everything is built up from these particles we hope that the knowledge on the physics of the building blocks leads to the understanding of more complex systems. Experimentalists are looking for different particles and interactions, which are then put into a consistent framework by theorists, who have to solve the equations, too. The most popular experimental way to gain information is to collide a few particles and look what happened. Many more particles were participating in these high energy processes in the early Universe (Big Bang) or are participating in present day heavy ion experiments (Little Bang). Note, that for both cases the baryonic densities are much smaller than the typical hadronic scales (and can be treated as zero).

At temperatures around $T \approx 200$ MeV a transition happened, which is related to the spontaneous breaking of the chiral symmetry of QCD. The nature of the QCD transition affects our understanding of the Universe's evolution (see ref.[1] and references therein). In a strong first-order phase transition the quark–gluon plasma supercools before droplets of hadron gas are formed. These droplets grow, collide and merge, during which gravitational waves could be produced[2]. Baryon-enriched nuggets could remain between the bubbles, contributing to dark matter. The hadronic phase is the initial condition for nucleosynthesis, so inhomogeneities in this phase could have a strong effect on nucleosynthesis. As the first-order phase transition weakens, these effects become less pronounced. Since about 99% of the mass of the visible Universe is generated during this transition, it is of extreme importance to understand its details. As we will see our calculations provide strong evidence that the QCD transition is a cross-over and thus the above scenarios – and many others – are ruled out. In addition we determine the absolute scale of this transition in physical units. The determination of the absolute scale pins down the temperature and time scale in the early Universe and has a huge impact on present and future heavy ion experiments.

Quantum chromodynamics (QCD) is the theory of the strong interaction, explaining (for example) the binding of three almost massless quarks into a much heavier proton or neutron – and thus most of the mass of the visible Universe. The strong interaction is

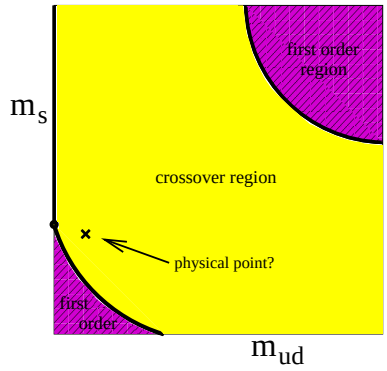


Figure 1. The phase diagram of QCD on the hypothetical light quark mass versus strange quark mass plane. Thick lines correspond to second order phase transitions, the purple regions represent first order phase transitions and the yellow region represents an analytic cross-over.

responsible for the energy producing processes in the sun or in nuclear reactors. The standard model of particle physics predicts a QCD-related transition. At low T, the dominant degrees of freedom are colourless bound states of hadrons (such as protons and pions). However, QCD is asymptotically free, meaning that at high energies or temperatures the interaction gets weaker and weaker[3,4], causing hadrons to break up. This behaviour underlies the predicted cosmological transition between the low T hadronic phase and a high T quark–gluon plasma phase (for simplicity, we use the word 'phase' to characterize regions with different dominant degrees of freedom). Despite enormous theoretical effort, the nature of this $T > 0$ QCD transition (that is, first-order, second-order or analytic cross-over) remained ambiguous. The reason for that is the extreme complication when one tries to solve the equations of QCD. Since they describe the strong interaction they are strongly coupled equations, which seem to be impossible to be solved analytically. The only known systematic technique which could give a final answer is lattice QCD.

QCD is a generalised version of quantum electrodynamics (QED). The Euclidean Lagrangian with gauge coupling g and with a quark mass of m can be written as $\mathcal{L}$=$-1/(2g^2)$Tr$F_{\mu\nu}F_{\mu\nu}$+$\bar{\psi}\gamma_\mu(\partial_\mu$+$A_\mu$+m$)\psi$, where $F_{\mu\nu}$=$\partial_\mu A_\nu$-$\partial_\nu A_\mu$+$[A_\mu,A_\nu]$. In QED the gauge field A_μ is a simple real field, whereas in QCD it is a 3×3 matrix. Consequently the commutator in $F_{\mu\nu}$ vanishes for QED, but it does not vanish in QCD. The ψ fields also have an additional "colour" index in QCD, which runs from 1 to 3. Different types of quarks are represented by fermionic fields with different m. The action S is defined as the four-volume integral of $\mathcal{L}$. The basic quantity we determine is the the partition function Z, which is the sum of the Boltzmann factors $\exp(-S)$ for all field configurations. Partial derivatives of Z with respect to m give rise to the order parameters we studied here.

Lattice QCD (c.f.[5]) discretises the above Lagrangian on a four-dimensional lattice and extrapolates the results to vanishing lattice spacing ($a \longrightarrow 0$). A convenient way to carry out this discretisation is to put the fermionic variables on the sites of the lattice, whereas the gauge fields are treated as 3×3 matrices connecting these sites. In this sense, lattice QCD is a classical four-dimensional statistical physics system. One important difference compared to three dimensional systems is that T is determined by the additional, Euclidean

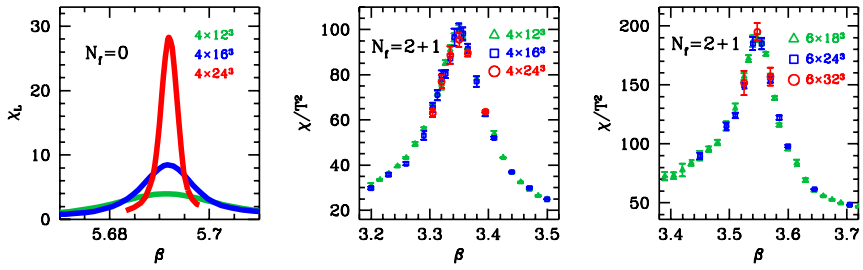


Figure 2. The V dependence of the susceptibility peaks for pure SU(3) gauge theory (Polyakov susceptibility, left panel) and for full QCD (chiral susceptibility on N_t= 4 and 6 lattices, middle and right panels, respectively).

time extension (N_t): T=1/($N_t a$). Keeping T fixed (such as T_c, the transition temperature) one can reduce a and approach the continuum limit by increasing N_t.

2 The Nature of the QCD Transition

The standard picture for the QCD phase diagram on the light quark mass (m_{ud}) versus strange quark mass (m_s) plane is shown by Fig. 1. It contains two regions at small and at large quark masses, for which the $T > 0$ QCD transition is of first order. Between them one finds a cross-over region, for which the $T > 0$ QCD transition is an analytic one. The first order transition regions and the cross-over region are separated by lines, which correspond to second order phase transitions. The location of the physical point (thus the nature of the QCD transition) was a priori unknown.

There are lattice results for the nature of the transition (for the two most popular lattice formulations see refs[6,7]), although they have unknown systematical errors. We emphasize that from the lattice point of view two "ingredients" are necessary to eliminate the systematical errors.

The first ingredient is to use physical quark masses. Owing to the computational costs this is a great challenge in lattice QCD. Previous analyses used computationally less demanding non-physically large quark masses. However, these choices have limited relevance. The order of the transition depends on the quark mass. For example, in three-flavour QCD for vanishing quark masses the transition is of first-order. For intermediate masses it is a cross-over. For infinitely heavy quark masses the transition is again first-order. For questions concerning the restoration of chiral symmetry (such as the order of the transition), a controlled extrapolation from larger quark masses (such as chiral perturbation theory) is unavailable, and so the physical quark masses should be used directly.

The second ingredient is to remove the uncertainty associated with the lattice discretization. These errors disappear in the continuum limit; however, they strongly influence the results at non-vanishing lattice spacing. E.g. in three-flavour unimproved staggered QCD, using a lattice spacing of about 0.28 fm, the first-order and the cross-over regions are separated by a pseudoscalar mass of $m_{\pi,c}$≈300 MeV. Studying the same three-flavour theory with the same lattice spacing, but with an improved p4 action (which has different

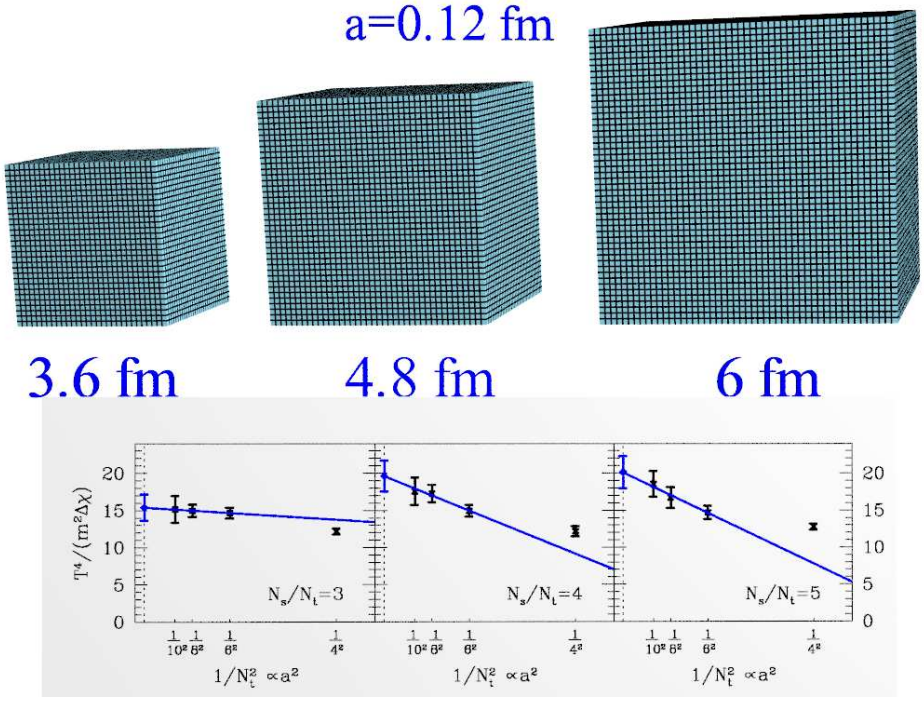


Figure 3. The upper pictures show three different physical V-s with our finest discretization. The lower plots show the dimensionless quantity $T^4/(m^2\Delta\chi)$ as a function of a^2 and their continuum extrapolated values. No V dependence is observed.

discretization errors) we obtain $m_{\pi,c}\approx$70 MeV. In the first approximation, a pseudoscalar mass of 140 MeV (which corresponds to the numerical value of the physical pion mass) would be in the first-order transition region, whereas using the second approximation, it would be in the cross-over region. The different discretisation uncertainties are solely responsible for these qualitatively different results[8]. Therefore, the proper approach is to use physical quark masses, and to extrapolate to vanishing lattice spacings.

Our work eliminates both of the above uncertainties[9].

We will study the finite size scaling of the lattice chiral susceptibilities $\chi(N_s, N_t)$=$\partial^2/(\partial m_{ud}^2)(T/V)\cdot \log Z$, where m_{ud} is the mass of the light u,d quarks and N_s is the spatial extension. This susceptibility shows a pronounced peak around T_c. For a real phase transition the height of the susceptibility peak increases and the width of the peak decreases when we increase the volume (V). For a first-order phase transition the finite size scaling is determined by the geometric dimension, the height is proportional to V, and the width is proportional to $1/V$. Such an example can be seen on the left panel of Fig. 2. In the pure SU(3) gauge theory –QCD with no fermions– the transition is known to be of first order. Thus, the characteristic increase of the analogous susceptibility can be seen. For a second-order transition the singular behaviour is given by some power of V, defined by the critical exponents. The picture would be completely different for an analytic

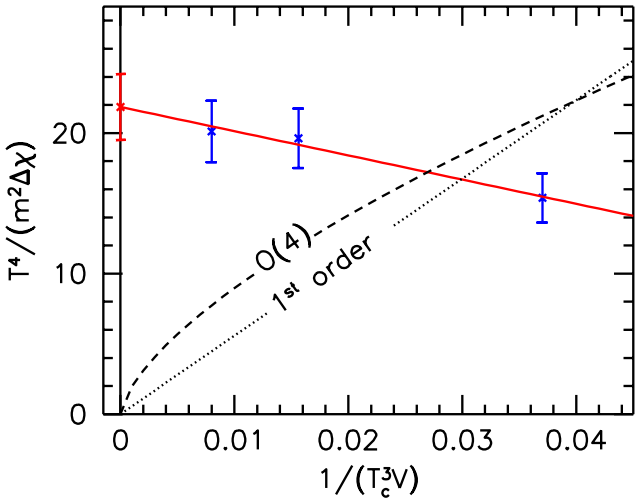


Figure 4. Continuum extrapolated susceptibilities $T^4/(m^2\Delta\chi)$ as a function of $1/(T_c^3 V)$. For true phase transitions the infinite V extrapolation should be consistent with zero, whereas for an analytic cross-over the infinite V extrapolation gives a non-vanishing value. The continuum-extrapolated susceptibilities show no phase-transition-like volume dependence, though V changes by a factor of five. The V$\rightarrow\infty$ extrapolated value is 22(2) which is 11σ away from zero. For illustration, we fit the expected asymptotic behaviour for first-order and O(4) (second order) phase transitions shown by dotted and dashed lines, which results in chance probabilities of 10^{-19} (7×10^{-13}), respectively.

cross-over. There would be no singular behaviour and the susceptibility peak does not get sharper when we increase V; instead, its height and width will be V independent for large V. Quite interestingly this behaviour is observed for full QCD, the susceptibility peaks are almost V independent (see the middle and right panels of Fig. 2 for the susceptibilities for the light quarks for $N_t = 4$ and 6, for which we used aspect ratios $r = N_s/N_t$ ranging from 3 to 6 and 3 to 5, respectively).

Unfortunately, these curves do not say much about the continuum behaviour of the theory. In principle a phenomenon as unfortunate as that in the three-flavour theory could occur[8], in which the reduction of the discretization effects changed the nature of the transition for a pseudoscalar mass of $\approx$140 MeV.

In order to clarify this issue (which is a genuine $T > 0$ effect) we subtract the T= 0 susceptibility and study only the difference between $T \neq 0$ and T= 0 at different lattice spacings. This leads to $m^2\Delta\chi$, which we study. To give a continuum result for the nature of the transition we carry out a finite size scaling analysis of the dimensionless quantity $T^4/(m^2\Delta\chi)$ directly in the continuum limit. For this study we need the height of the susceptibility peaks in the continuum limit for fixed physical V. The continuum extrapolations are done using four different lattice spacings. The V-s at different lattice spacings are fixed in units of T_c, and thus VT_c^3=3^3,4^3 and 5^3 were chosen. Fig. 3 shows one typical discretization of our different physical sizes and the continuum extrapolation for these three volumes. N_t=4 results are off, N_t=6,8 and 10 show a good $a^2 \propto 1/N_t^2$ scaling.

Having obtained the continuum values for $T^4/(m^2\Delta\chi)$ at fixed physical V-s, we study the finite V scaling of the results. Fig. 4 shows our final results. The V dependence shows

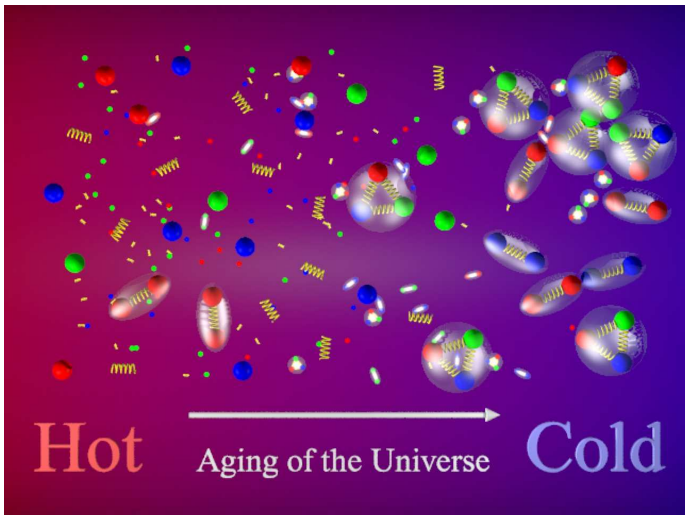


Figure 5. A schematic view of the transition between the quark-gluon plasma and the hadronic phase. Free quarks and gluons are confined to hadrons as the Universe cools. The transition is smooth, no singularity appears.

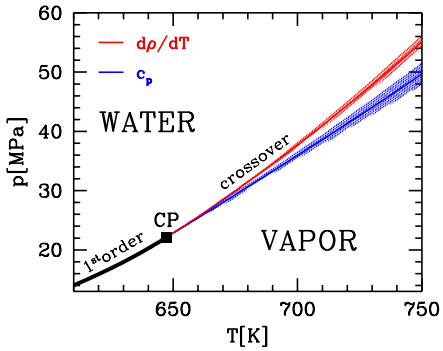


Figure 6. The water-vapour phase diagram.

that there is no true phase transition but only an analytic cross-over in QCD.

Thus, no droplet formation happened, the quark-gluon plasma went through smoothly to a hadronic phase. This smooth transition is shown as a cartoon on Fig. 5.

3 The Transition Temperature

An analytic cross-over, like the QCD transition has no unique T_c. A particularly nice example for that is the water-vapour transition (c.f. Fig. 6). Up to about 650 K the transition is a first order one, which ends at a second order critical point. For a first or second order phase transition the different observables (such as density or heat capacity) have their singularity (a jump or an infinitely high peak) at the same pressure. However, at even

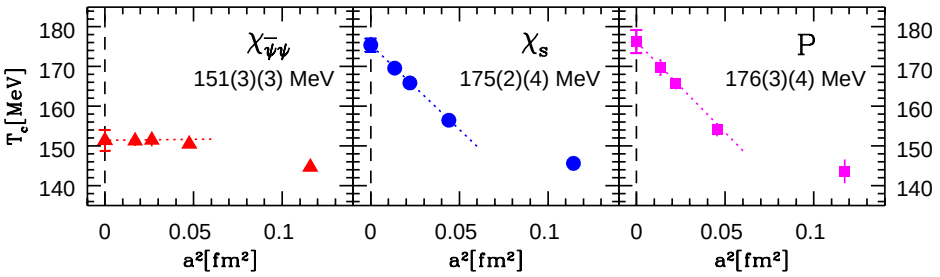


Figure 7. Continuum limits of T_c obtained from the renormalized chiral susceptibility ($m^2\Delta\chi_{\bar{\psi}\psi}/T^4$), strange quark number susceptibility (χ_s/T^2) and renormalized Polyakov-loop (P_R).

higher T-s the transition is an analytic cross-over, for which the most singular points are different. The blue curve shows the peak of the heat capacity and the red one the inflection point of the density. Clearly, these T_c-s are different, which is a characteristic feature of an analytic transition (cross-over). In QCD we will study the chiral and the quark number susceptibilities and the Polyakov loop. Usually they give different T_c values, but there is nothing wrong with it. As was illustrated by the water-vapour transition it is a physical ambiguity, related to the analytic behaviour of the transition. There is another, non-physical, ambiguity. If we used different observables (particularly at large lattice spacings) to set the lattice spacings we obtain different overall scales. They lead to different T_c values. This ambiguity disappears in the continuum limit.

According to our experiences, at finite lattice spacings, the best choice is the kaon decay constant f_k. It is known experimentally (in contrast to the characteristics of the static potential), thus no intermediate calculation with unknown systematic errors is involved. Furthermore, it can be measured on the lattice quite precisely. The continuum extrapolated T_c values obtained[10] for different quantities are shown on Fig. 7.

There is a surprising several sigma effect. The remnant of the chiral transition happens at a quite different T than that of the deconfining transition. It is a robust effect, since the Polyakov transition region is quite off the χ-peak, and the χ-peak is far from the inflection point of the Polyakov loop. This quite large difference is also related to the fact that the transition is fairly broad. The widths are around 30-40 MeV.

One can set the overall scale by r_0 (it is defined by the static potential between a quark and an antiquark $dV/dr^2 \cdot r_0^2$=1.65). On coarse lattices different choices might lead to ambiguities for the T_c, which is illustrated for our data on Fig. 8. Using only N_t=4 and 6 the continuum extrapolated T_c-s are quite different if one took r_0 or f_K to determine the overall scale. This inconsistency indicates, that these lattice spacings are not yet in the scaling region (a similar ambiguity is obtained by using the p4 action of[11]). Having N_t=4,6,8 and 10 results this ambiguity disappears (as usual N_t=4 is off), these lattice spacings are already in the scaling region (at least within our accuracy). This phenomenon is not surprising. At large a-s the scale cannot be determined unambiguously. This underlines the importance of the continuum limit we carried out.

The ambiguity related to the inconsistent continuum limit is unphysical, and it is resolved as we approach the continuum limit (c.f. Fig. 8). The differences between the T_c

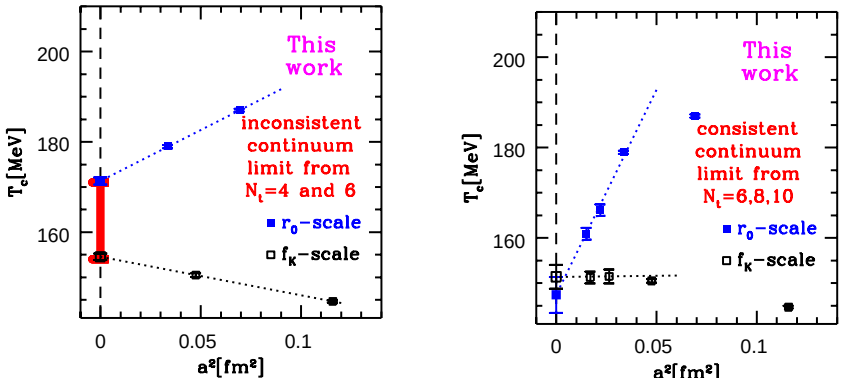


Figure 8. Continuum extrapolations based on N_t=4 and 6 (left panel: inconsistent continuum limit) and using N_t=6,8 and 10 (right panel: consistent continuum limit).

values for different observables are physical, this is a consequence of the cross-over nature of the QCD transition.

4 Summary

The nature of the QCD transition was determined, which turned out to be an analytic cross-over. This result excludes most cosmic relic scenarios. The transition temperature was determined for several observables. Since the transition is a broad one, there is a 30-40 MeV difference between the T_c values defined by different observables.

Acknowledgments

A significant part of the calculations were carried out on the BlueGene/L machine in Jülich.

References

1. D. J. Schwarz, Annalen Phys. **12** (2003) 220 [arXiv:astro-ph/0303574].
2. E. Witten, Phys. Rev. D **30**, 272 (1984).
3. D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973).
4. H. D. Politzer, Phys. Rev. Lett. **30** (1973) 1346.
5. I. Montvay and G. Munster, "Quantum fields on a lattice," *Cambridge, UK: Univ. Pr. (1994) 491 p.*
6. F. R. Brown *et al.*, Phys. Rev. Lett. **65**, 2491 (1990).
7. A. Ali Khan *et al.* [CP-PACS Collaboration], Phys. Rev. D **63**, 034502 (2001) [arXiv:hep-lat/0008011].
8. F. Karsch et al., Nucl. Phys. Proc. Suppl. **129** (2004) 614 [arXiv:hep-lat/0309116].
9. Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz and K. K. Szabo, Nature **443** (2006) 675.
10. Y. Aoki, Z. Fodor, S. D. Katz and K. K. Szabo, Phys. Lett. B **643** (2006) 46.
11. M. Cheng et al., Phys. Rev. D **74**, 054507 (2006) [arXiv:hep-lat/0608013].

Astrophysics

Peter L. Biermann

Max Planck Institut für Radioastronomie, Bonn, Germany
E-mail: plbiermann@mpifr-bonn.mpg.de

Dept of Physics and Astronomy, University of Bonn, Germany

Dept of Physics and Astronomy, University of Alabama, Tuscaloosa, AL, USA

Dept of Physics, University of Alabama at Huntsville, AL, USA

New problems in fundamental physics can be tackled with the availability of yet faster and bigger supercomputers. As it has become clear recently, that the particles of the highest energy observed in the cosmos come from active galactic nuclei (Auger-Coll. 2007), always found in the centre of galaxies, and embedded in gas which has been heavily enriched by supernova explosions, our probes to understand the deep structure of matter now require a physical understanding of all the processes leading to the activity. Here we highlight a) the evolution of galaxies in clusters, their stripping, and their energy input from active galactic nuclei, b) the explosion of massive stars to yield supernovae and occasionally black holes, and c) the evolution of binary black holes and active galactic nuclei.

Brüggen & Roediger discuss new simulations to use clusters of galaxies as probes to understand the *evolution of galaxies*: This involves the evolution of the stellar population and the gas in galaxies, which is enriched by supernova explosions; these galaxies can then be partially stripped as they move through the cluster. While such calculations have been done for some thirty years, these new simulations achieve a new level of sophistication. A question is what happens when clusters merge, and most of the intra-cluster gas slops around, as then the gas distribution itself will be hugely temporally variable, and will strip galaxies far beyond its normal sphere of influence. The authors also simulate episodes of the activity of a central massive black hole, and its concurrent heating of the intra-cluster gas via weak shock-waves; as the heating transport does not disturb the abundance gradient in the cluster, it is an unsolved question, how the heating actually happens. What is an interesting question is whether we can use this heating as a quantitative path to understand the energy transport of jets better in cases, when the jet is driven by the spin-down of a black hole (Blandford & Znajek 1977, 1978), and we miss any other method to quantify the jet's properties.

The quest of why *massive stars explode*, and what the physical mechanism is, has bedeviled us for decades. Even today we still think of new ideas, how Nature might do this. 1964 the concept was born that rotation and magnetic fields might turn a collapse into an explosion, and only this decade this idea was turned into successful simulations by Bisnovatyi-Kogan and Moiseenko. 1966 Colgate and White suggested that neutrinos cause the explosion, an idea which has found a lot of resonance in the community. Janka, Marek & Müller describe the latest simulations following this idea. Clearly, the combination of high magnetic fields, high rotation, and full 3D treatment is still a challenge. As some expect the mechanism for these explosions to also produce Gamma Ray Bursts – believed

to be the signature of the formation of black holes at the stellar mass scale – under special circumstances, any further evolution of such simulations will yield great benefits.

Almost all galaxies harbor a *massive black hole at their centre*, and one primary mechanism of galaxies to grow and evolve is by mergers. Consequently, it can be inferred, that also black holes grow by mergers; as usually these mergers involve a mixing of orbital and intrinsic spins, this implies the judicious application of Einstein's theory of General Relativity. On the other hand, accretion of gaseous matter or entire stars can fuel the observed activity of black holes, visible as relativistic jets and ultra-violet emitting accretion disks, the most energetic observed phenomena in the universe. We now have direct evidence that such activity is the fundamental cause to accelerate particles to the highest energy measured yet in the universe, and so with such simulations as described by Spurzem et al. here, we will be able to ultimately turn these super-massive black holes and their phenomena into tools to understand the deep structure of matter.

The formation and evolution of galaxies, their gas, stars and black hole population, as well as their massive central black holes, and all the concurrent physical phenomena can ultimately be expected to yield deep insight into the *structure of matter*. One of the strongest tools in this endeavour is the careful use of super-computer simulations.

Acknowledgments

Support for PLB is coming from the AUGER membership and theory grant 05 CU 5PD 1/2 via DESY/BMBF.

References

1. Correlation of the highest energy cosmic rays with nearby extragalactic objects, AUGER-collaboration (includ. P.L. Biermann), *Science*, **318**, 939 - 943 (9 November 2007); arXiv:0711.2256.
2. Magnetorotational supernovae, Ardeljan, N. V., Bisnovatyi-Kogan, G. S, Moiseenko, S. G., *Month. Not. Roy. Astr. Soc.*, **359**, p. 333 - 344 (2005).
3. The Explosion of a Rotating Star As a Supernova Mechanism, Bisnovatyi-Kogan, G.S., *Astron.Zh. (Sov. Astron.)*, **47**, 813 (1970).
4. The Hydrodynamic Behavior of Supernovae Explosions, Colgate, S. A., White, R. H., *Astrophys. J.*, **143**, 626 (1966).
5. Magnetic Collapse and the Nature of Intense Sources of Cosmic Radio-Frequency Emission, Kardashev, N. S., *Astron. Zh. (Sov. Astron.)*, **41**, 807 (1964).

Simulating Galaxies in Clusters

Marcus Brüggen and Elke Roediger

Jacobs University Bremen
Campus Ring 7, 28759 Bremen, Germany
E-mail: {*m.brueggen, e.roediger*}*@jacobs-university.de*

How galaxies form has puzzled astrophysics for decades and is among the key questions in contemporary astrophysics. Clusters of galaxies are good laboratories for the study of galaxy evolution. Among the processes that affect the life of a galaxy are the headwinds that the galaxies experience as they fly through the hot and dilute intracluster medium and the energetic outbursts from the supermassive black holes that sit in the centre of galaxies. Using the JUMP supercomputer, we have simulated these processes using adaptive-mesh hydrodynamical simulations.

1 Introduction

Galaxies are the building blocks of the universe. They are composed of stars, the interstellar medium (ISM) and dark matter halos. They show a variety of morphologies, ranging from ellipticals over spirals to irregulars, from dwarfs to giants. The galaxies themselves are part of still larger structures, they are members of groups and clusters of galaxies. A few galaxies are also isolated. It is a well-established fact that the evolution of galaxies is influenced by their environment. Moreover, the growth of galaxies is believed to be strongly determined by the supermassive black hole that resides in the centre of every galaxy, which we will discuss in Sec. 4.

Especially the properties of disk galaxies change with environment: In denser regions, the galaxies tend to contain less neutral gas, show a weaker star formation activity and redder colours than galaxies in sparse regions (e.g. Van Gorkom (2004)[19], Goto (2003)[3]). Several processes have been proposed to explain these features. One of the most important processes is ram pressure stripping (RPS) which works as follows: Besides galaxies, clusters also contain a large amount of rarefied gas – the intra-cluster medium (ICM). In fact, there is more mass in this dilute intra-cluster gas than in all the stars of the cluster galaxies. As galaxies move through a cluster, they also move through the ICM. The ram pressure caused by these motions can push out (parts of) their gas disks.

A galaxy's gas is its raw material for star formation, therefore such ram pressure stripping has severe consequences. E.g. the characteristic blue colour of spiral galaxies is due to shortlived massive stars. When the star formation is reduced because the galaxy has lost its gas and no new stars are formed, the galaxy's colour changes from blue to red. So the process of RPS seems to be a good candidate to explain several of the differences between galaxies in different environments.

Moreover, several individual galaxies are known to show characteristics of RPS, i.e. an undisturbed stellar disk but a distorted and truncated gaseous disk (e.g. NGC 4522[5–7], NGC 4548[20]). Recently, observations of tails of stripped gas[9, 18] have been presented. They reveal long (up to 100 kpc) and narrow (10-20 kpc) gas tails that stretch away to one side of the galaxy. The structure of these tails differs from flaring, S-shaped to very narrow and nearly structureless. These observations provide an excellent opportunity to compare

to simulations and thus understand RPS in more detail. The stripped gas plays an important role in the chemical evolution of the ICM. Practically all metals (in astrophysics this means all elements heavier than hydrogen and helium) are produced in stars inside galaxies. Hence, the metals found in the ICM must originate from the cluster galaxies. Obviously, galactic gas lost by ram pressure stripping is a metal source for the ICM. Hydrodynamical simulations of the cluster crossing of a disk galaxy, i.e. where the galaxy is exposed to a varying ram pressure, had not been performed so far. All hydrodynamical simulations up to now[10, 17, 16, 13, 14] used a constant ram pressure. Additionally, the work of Schulz (2001)[17], Roediger (2005)[16], Roediger (2005)[13] has shown that the gas removal from the galactic potential does not happen instantaneously, but that it takes some time until the gas is accelerated enough to be unbound from the galaxy's potential. Thus, in cases of short ram pressure peaks, this time delay may play an important role. We have performed simulations where the galaxy is exposed to a varying ram pressure.

2 Method

We performed our simulations in three dimensions in Cartesian geometry using a modified version of the adaptive mesh refinement code FLASH (version 2.5). The FLASH code was developed by the Department of Energy-supported ASCI/Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago. FLASH is a modular block-structured AMR code, parallelised using the Message Passing Interface (MPI) library. It solves the Riemann problem on a Cartesian grid using the Piecewise-Parabolic Method (PPM). More details can be found under `http://flash.uchicago.edu`.

We chose a simulation box large enough to contain a substantial part of a whole galaxy cluster. The model galaxy is moving on a realistic orbit through the cluster, i.e. it is moving through the grid. Here, not only the strength of the ICM wind varies, but also its direction. For demonstration, Fig. 1 displays a slice through the simulation grid in the orbital plane of the galaxy.

The galaxy's orbit is marked by the black line. The position and size of the third frame (lhs column) in Fig. 2 is marked by the white rectangle. The initial diameter of the galaxy is about half the size of the white rectangle. Our model galaxy is a massive disk galaxy with a flat rotation curve at $200\,\mathrm{km\,s^{-1}}$. We studied different galaxy orbits in different clusters. The simulations need to cover an enormous range of scales (compare Figs. 1 and 2): The simulation box has to contain the whole orbit of the galaxy, which requires a box size of $(-1\,\mathrm{Mpc}, 1\,\mathrm{Mpc}) \times (-2\,\mathrm{Mpc}, 2\,\mathrm{Mpc}) \times (-1\,\mathrm{Mpc}, 1\,\mathrm{Mpc})$. On the other hand, the dynamics inside and close to the galaxy have to be resolved. At minimum, a resolution of $0.5\,\mathrm{kpc}$ is needed near the galaxy. Formally, this results in an effective resolution of $4096 \times 8192 \times 4096$ grid cells. Thanks to the adaptive mesh of the FLASH code and the supercomputers at the FZ Juelich, this is possible. A complete cluster crossing requires a physical runtime of $3\,\mathrm{Gyr}$. The typical runtime for these simulations is 2000-3000 CPUh, each run needs 40 GB of disk space and 30-40 GB memory. We are also running one simulation where the resolution is a factor of two better, which takes approximately 20 000 CPUh.

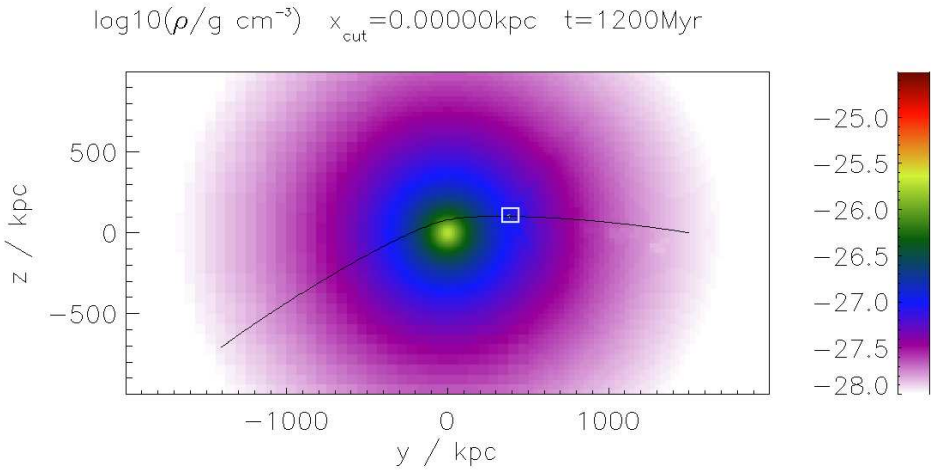


Figure 1. Density in the orbital plane of the galaxy. The orbit of the galaxy is shown by the black line. The white rectangle marks the position of the third frame in the left column in Fig. 2.

3 Results

As an example, Fig. 2 shows a series of snapshots of the density in the orbital plane of the galaxy. In the left column, just the vicinity of the galaxy is shown (compare to Fig. 1). The right column shows the distribution of stripped gas along the galaxy's orbit. A movie can be found here[a]. First, the galaxy moves subsonically, then it becomes supersonic. The gas disk becomes smaller and smaller as the galaxy is ram pressure stripped. In the last timestep shown the galaxy is already stripped completely, the peak in density is not located at the galactic centre anymore. The stripped gas is distributed all along the orbit.

There are several different aspects that can be investigated with these simulations:

3.1 Comparison to the Analytical Estimate

Ram pressure stripping is not only important for the stripped galaxies, but also for the ICM in the galaxy cluster. The galactic gas is enriched with heavier elements ("metals" in astrophysical terminology), and as the galaxies lose their gas, also the metals are distributed through the cluster. Simulations that aim at modelling the enrichment history of the ICM have to model the evolution of a whole galaxy cluster, including ICM, dark matter, and galaxies. This alone is computationally expensive, thus, they rely on analytical approximations for the single enrichment processes. For RPS, the analytical estimate explained below is commonly used.

The usual way to estimate the amount of gas loss for galaxies moving face-on follows the suggestion of Gunn&Gott (1972)[4]. Here, one compares the gravitational restoring force

[a] http://www.faculty.iu-bremen.de/eroediger/PLOTS/RPS_CLUSTER/dens_metdens_movie_x0.mpg

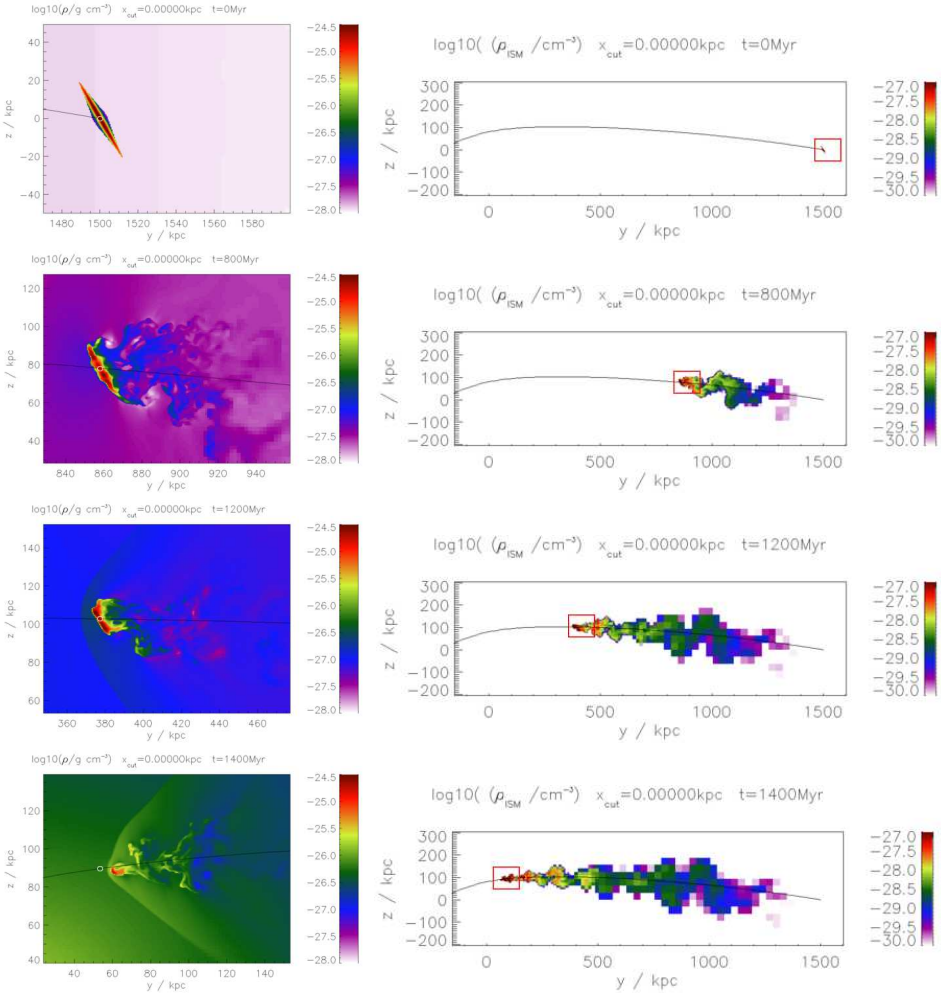


Figure 2. Time series of snapshots of a simulation run: The left column shows the gas density in the orbital plane. The small white circle marks the position of the galactic centre. In the second frame the galaxy moves still subsonic, then it becomes supersonic as indicated by the bow shock. In the last frame, the galaxy is already stripped completely. The frames in the right column show the density of the *galactic* gas only, i.e. they demonstrate the distribution of the stripped gas.
In all frames, the black line marks the galaxy's orbit. The frames in the left column show a much smaller part of the simulation grid. For orientation, the position and size of the lhs frames is marked by red rectangles in the rhs frames. However, even the frames in the right column do not show the complete simulation box (compare to Fig. 1).

per unit area and the ram pressure for each radius of the galaxy. At radii where the restoring force is larger, the gas can be retained, at radii where the ram pressure is larger, the gas will be stripped. The transition radius is called the stripping radius. Simulations with constant ram pressure generally found a good agreement between the analytical estimate and the

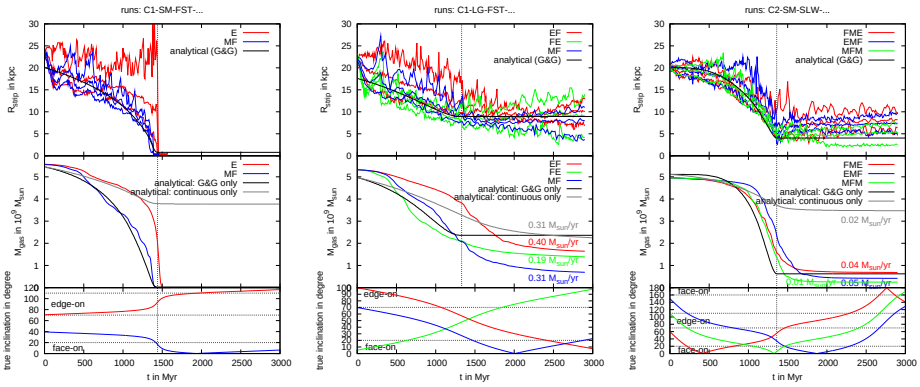


Figure 3. Comparison between analytical and numerical stripping radius (top panels) and bound gas mass (middle panels). The bottom panels display the evolution of the true inclination, i.e. the angle between the galaxy's rotation axis and the direction of motion. Each column is for one orbit (see title of column), different inclinations are colour-coded (see legend).
For the numerical stripping radius, the mean value (thick lines) as well as maximum and minimum radius (thin lines) as described in Roediger et al. (2006)[13] are shown.
Analytical estimates according to the Gunn&Gott criterion are shown as black lines for the stripping radius as well as for the remaining gas mass. For an explanation of continuous stripping see Roediger et al. (2006)[13] and Roediger & Brüggen (2007)[11].

simulated gas disk radius and mass. Although this estimate is derived for face-on galaxies, it holds well as long as the galaxy is not moving close to edge-on[13]. An aspect not captured by the analytical estimate is the delay in gas loss. Even if the ram pressure is large enough to remove some part from the galaxy, it takes some time until the gas is accelerated to the escape velocity.

In order to apply this estimate to our simulations, we use a time-dependent version of the classical Gunn&Gott criterion, i.e. for each timestep we compare the gravitational restoring force to the current ram pressure.

Figure 3 compares analytical estimates and numerical results for the stripping radius and the remaining gas mass for different runs. Each column is for one orbit, inside each column, different inclinations are colour-coded.

The analytical estimate is shown as black lines. This estimate neglects several aspects, e.g. inclination, a possible evolution of the galaxy (e.g. decrease in gas surface density) and the fact that the gas loss does not happen immediately, but with a certain delay. Despite these shortcomings, the analytical and numerical result for the stripping radius agree remarkably well. The only exception occurs on an orbit characterised by a narrow ram pressure peak of medium amplitude (rhs column of Fig. 3, the galaxy is expected to be stripped severely, but not completely), and here only during the ram pressure peak. If the galaxy moves face-on during the ram pressure peak, it is stripped completely, although the analytical estimate predicts a stripping radius of 4 kpc. If the galaxy moves with medium inclination through the cluster centre, it can retain a larger disk than predicted. However, also for this orbit, during the first Gyr of these simulations, the agreement between analytical estimate and numerical results is good. The differences are found only during the ram pressure peak. For the galaxies with medium inclination, this behaviour is caused by the delay of the gas loss combined with the shortness of the ram pressure peak.

3.2 Distribution of Stripped Gas Along the Orbit:

The rhs column of Fig. 2 demonstrates that the galaxy is losing gas continuously all along the orbit. The knowledge about where exactly the galaxy deposits which amount of stripped gas is crucial for the investigation of the metal enrichment history of the ICM. An analytical estimate for the ISM distribution along the orbit can be derived from the analytical estimate of the stripping radius explained in Sect. 3.1. According to this estimate, the galaxy does not lose any gas after pericentre passage, all stripped gas should be deposited before the pericentre. However, the simulations show that the galaxy loses gas after pericentre passage, and the galaxy loses gas more slowly than predicted. Thus, the stripped gas is spread much wider along the orbit. Details of the distribution depend also strongly on inclination.

4 Feedback by Active Galactic Nuclei

Recent observations of galaxy clusters show a multitude of physical effects that occur when powerful jets launched by supermassive black holes interact with the surrounding medium (See Brüggen & Kaiser 2002[1]). While these effects are widely believed to be crucial for the formation of structure in the universe, they are still poorly understood. Clusters of galaxies are excellent laboratories for studying the interaction between active galactic nuclei (AGN) and diffuse gas. Recent observational evidence demonstrates that the lives of AGN and their environment are closely intertwined. This complex pattern of processes has been simulated with unprecedented realism by our group at the Jacobs University Bremen using the JUMP supercomputer.

Again we used a modified version of the adaptive-mesh hydrodynamics code FLASH described above for these simulations. It uses a criterion based on the dimensionless second derivative of a fluid variable to refine or derefine the grid. Collisionless matter, i.e. stars and dark matter, were represented by particles that interact gravitationally with each other and the gas. Our initial models were extracted from cosmological smoothed-particle hydrodynamics simulations and typically included around 700,000 dark matter and star particles. We modified the FLASH code to follow the central black hole and modelled the jet as inflow boundary conditions that lie within the computational domain. Moreover, we developed a fast multigrid gravity solver in order to be able to simulate the galaxy cluster for more than 300 million years.

We have achieved the following:

- We have studied the interaction of the jet with its environment, for the first time taking into account the dynamic nature of the cluster gas. The simulations successfully reproduce the observed morphologies of radio sources in clusters. We find that cluster inhomogeneities and large-scale flows have significant impact on the morphology of the radio source.

- The ICM has a metallicity of about 1/3 solar. However, cooling core clusters, i.e. clusters with a centrally peaked X-ray brightness, show peaked abundance profiles. A number of observations indicates that supernovae in the central galaxy are mainly responsible for the metal enrichment in the central part of clusters. However, the

observed metallicity profiles are much broader than the light profiles of the central cluster galaxy. Hence, the difference in the light and the metal distributions are interpreted as the result of transport processes that have mixed metals into the ICM. While it appears to be established that the metals produced by the central galaxy are dispersed into the ICM to form the broad abundance peaks, it remains unclear what the mechanism is via which the metals are transported. As one likely mechanism, we studied the effect of AGN-inflated bubbles that rise buoyantly and lift metal-rich gas upwards. We demonstrated that AGN can account for the distribution of metals in clusters of galaxies (for more details see Roediger et al. (2007)[12])

- Using realistic 3D simulations of jets in a galaxy cluster, we address the question what fraction of AGN energy is dissipated in shocks. We find that weak shocks that encompass the AGN have Mach numbers of 1.1-1.2 and dissipate at least 2 per cent of the mechanical luminosity of the AGN. In most cases this is sufficient to balance the radiative losses of the gas. In a realistic cluster medium, even a continuous jet can lead to multiple shock structures, which may lead to an overestimate of the AGN duty cycles inferred from the spatial distribution of waves. (for more details see Brüggen et al. (2007)[2])

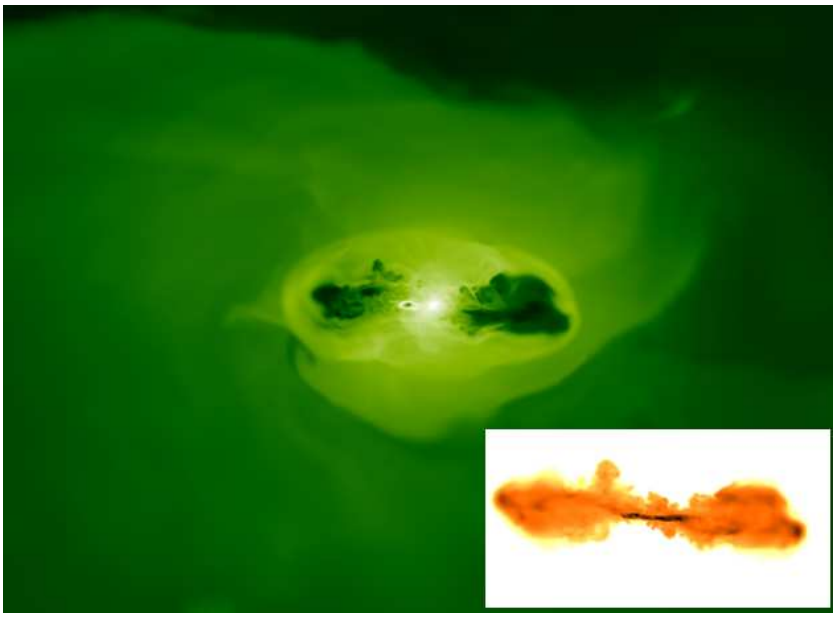


Figure 4. Snapshots of the gas density in a simulation of a jet in a cluster of galaxies. One can see how the jet drives shocks into the intracluster medium. The inset shows the simulated radio emission from this jet. The shock surrounding the two cavities are also clearly visible.

5 Concluding Remarks

The simulation of the complex multi-physics and multi-scale processes in galaxies is just at its beginning. It is a formidable computational challenge that will continue to drive the requirements for supercomputing infrastructure. Here we have presented some recent successes using grid-based simulations on adaptive meshes.

Acknowledgments

MB and ER acknowledge the support by the DFG grant BR 2026/3 within the Priority Programme "Witnesses of Cosmic History" and supercomputing grants at the John von Neumann Institute for Computing in Jülich. The results presented were produced using the FLASH code, a product of the DOE ASC/Alliances-funded Center for Astrophysical Thermonuclear Flashes at the University of Chicago.

References

1. Brüggen, M., & Kaiser, C. R. 2002, Nature, 418, 301.
2. Brüggen, M., Heinz, S., Roediger, E., Ruszkowski, M., & Simionescu, A. 2007, MNRAS, 380, L67.
3. Goto, T., Yamauchi, C., Fujita, Y., et al. 2003, MNRAS, 346, 601.
4. Gunn, J. E. & Gott, J. R. 1972, ApJ, 176, 1.
5. Kenney, J. D. P. & Koopmann, R. A. 1999, AJ, 117, 181.
6. Kenney, J. D. P. & Koopmann, R. A. 2001, in ASP Conf. Ser., Vol. 240, Gas and Galaxy Evolution, ed. J. E. Hibbard, M. P. Rupen, & J. J. van Gorkom (San Francisco: ASP), 577.
7. Kenney, J. D. P., van Gorkom, J. H., & Vollmer, B. 2004, AJ, 127, 3361.
8. Koopmann, R. A., Haynes, M. P., & Catinella, B. 2006, AJ, 131, 716.
9. Oosterloo, T. & van Gorkom, J. 2005, A&A, 437, L19.
10. Quilis, V., Moore, B., & Bower, R. 2000, Science, 288, 1617.
11. Roediger, E., & Brüggen, M. 2007, MNRAS, 380, 1399.
12. Roediger, E., Brüggen, M., Rebusco, P., Böhringer, H., & Churazov, E. 2007, MNRAS, 375, 15.
13. Roediger, E. & Brüggen, M. 2006, MNRAS, 369, 567.
14. Roediger, E., Brüggen, M., & Hoeft, M. 2006, MNRAS, 371, 609.
15. Roediger, E., Brüggen, M., Rebusco, P., Böhringer, H., & Churazov, E. 2007, MNRAS, 375, 15.
16. Roediger, E. & Hensler, G. 2005, A&A, 433, 875.
17. Schulz, S. & Struck, C. 2001, MNRAS, 328, 185.
18. Sun, M., Jones, C., Forman, W., et al. 2006, ApJ, 637, L81.
19. van Gorkom, J. H. 2004, in Carnegie Observatories Astrophysics Series, Vol. 3, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution, ed. J. S. Mulchaey, A. Dressler, & A. Oemler (Cambridge: Cambridge University Press), 306.
20. Vollmer, B., Cayatte, V., Boselli, A., Balkowski, C., & Duschl, W. J. 1999, A&A, 349, 411.

Unravelling the Explosion Mechanism of Core-Collapse Supernovae

Hans-Thomas Janka, Andreas Marek, and Bernhard Müller

Max Planck Institute for Astrophysics
Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
E-mail: {thj, amarek, bjmuellr}@mpa-garching.mpg.de

We report here on current progress in understanding the cataclysmic death of massive stars, in course of which the stellar core collapses to a neutron star or black hole and the star itself gets disrupted in a gigantic supernova explosion. Improved numerical tools and the increasing power of modern supercomputers like those of the NIC allow us now to simulate with unprecedented sophistication and detailedness the extremely complex physical processes that take place in the deep interior of a dying star. This has led to the discovery of new phenomena in the supernova core, for example a generic hydrodynamic instability of the stagnant supernova shock against nonradial deformation and the excitation of gravity-wave activity in the nascent neutron star. Both can have a supportive effect on the inauguration of the explosion, the former by improving the conditions for energy deposition by neutrino heating in the postshock gas, the latter by supplying the developing explosion with a flux of acoustic energy that adds to the energy transferred by neutrinos. Here we show recent results of self-consistent, two-dimensional supernova explosion models computed by the Garching group for 9, 11, and 15 solar mass stars.

1 Supernova Explosions and Their Modeling

Supernova explosions mark the death of massive stars with a mass of roughly more than eight times that of the sun. They are among the most powerful and most dramatic events in the universe. While the inner, dense core of the evolved star becomes gravitationally unstable and within seconds collapses to a neutron star, and eventually sometimes to a black hole, most of the stellar gas is ejected with a kinetic energy that equals the radiation loss during the life of the sun. The hot debris of the disrupted star expands into the interstellar space with velocities of several thousand kilometers per second. It shines for weeks nearly as bright as a whole galaxy. This brilliant spectacle, however, is only a wimpy by-product of what happens at the centre of the supernova: The neutron star forming there emits hundred times more than the explosion energy in neutrinos, weakly interacting elementary particles that are abundantly produced in the hot and extremely dense matter of the nascent neutron star. These neutrinos carry away the enormous gravitational binding energy that is released when the moon-sized stellar core with one and a half times the mass of the sun collapses to the hundred times more compact supernova remnant.

Due to the huge energy radiated in neutrinos, these particles have long been speculated to be the driving agent of the stellar explosion. Colgate and White[1] in a seminal paper in 1966 not only proposed gravitational binding energy to be the primary energy source of core-collapse supernovae, but also that the intense flux of escaping neutrinos transfers the energy from the imploding core to the ejected stellar mantle. Nearly twenty years later, Bethe and Wilson[2] were the first who described in detail the way how this might happen, interpreting thereby the physics that played a role in hydrodynamic simulations performed by Wilson.

These pioneering computer simulations of the so-called delayed neutrino-driven explosion mechanism were still conducted in spherical symmetry. The mechanism turned out to be successful only when the neutron star was assumed to become a more luminous neutrino source by convection accelerating the energy transport out of the dense interior. The thus enhanced neutrino emission led to stronger neutrino heating in the overlying layers of the exploding star. Later theoretical studies and the first multi-dimensional computer models, which became available only in the mid 1990's, however, suggested that convection inside the neutron star does not have the necessary big effect. Instead, these simulations demonstrated that the neutrino-heated layers around the forming neutron star are unstable to vigorous convective overturn. This can raise the efficiency of the neutrino energy deposition and thus can have a supportive effect for the supernova explosion. The first such two-dimensional (i.e. axisymmetric) and three-dimensional simulations, however, suffered from a severe drawback: the essential physics of the neutrino transport and neutrino-matter interactions is so complex that it could be treated only in a grossly simplified way, namely by a so-called "grey diffusion approximation". This means that the energy-dependence of the neutrino interactions was ignored and the spatial propagation was approximated by assuming that neutrinos diffuse through the dense neutron star medium until they decouple at the "neutrinosphere" close to the surface of the compact remnant. The historical development of these theoretical studies of the supernova explosion mechanism is resumed in a recent review by Janka et al.[3].

Our project has the goal of advancing the modelling of stellar core collapse to the next level of sophistication. To this end, the Garching group developed a new neutrino transport code (see Rampp and Janka[4]) that is able to describe the propagation of neutrinos through the supernova interior with much higher accuracy than before. It fully accounts for the energy dependence of the problem and for the gradual transition from neutrino diffusion at high densities in the neutron star interior to free streaming of neutrinos in the much more dilute stellar layers far outside of the neutron star. Results of simulations with this latest generation of modelling tools will be presented in Section 3.

A better understanding of the explosion mechanism of core-collapse supernovae is crucial for a large variety of important problems in astrophysics and nuclear astrophysics, but has also very high relevance for fields like particle physics and gravitational physics. What are the properties of supernova explosions and how do they depend on the progenitor star that explodes? Which stars give birth to neutron stars and which ones to black holes when they die? What kind of radioactive elements are synthesized during the explosion and which amounts of them get ejected? Are supernovae the long-sought site of the formation of the most neutron-rich elements like gold, lead, thorium, and uranium in the so-called r-process? What is the signature of neutrino and gravitational-wave emission from collapsing stars, and what can these signals tell us about the physics in the deep interior of the supernova, in particular at the extreme and possibly exotic conditions in the forming neutron star? These and many other questions require a better theoretical insight into the processes that lead to the explosion. Because light is emitted from the surface of the disrupted star and its observation thus yields only indirect information of the events in the centre, the progress in supernova theory largely depends on the possibilities of computational modelling. The extraordinary complexity of the involved processes, which combine multi-dimensional hydrodynamics, relativistic gravity, neutrino and particle physics as well as nuclear physics, poses a major computational challenge. The availability of top-end computing platforms is

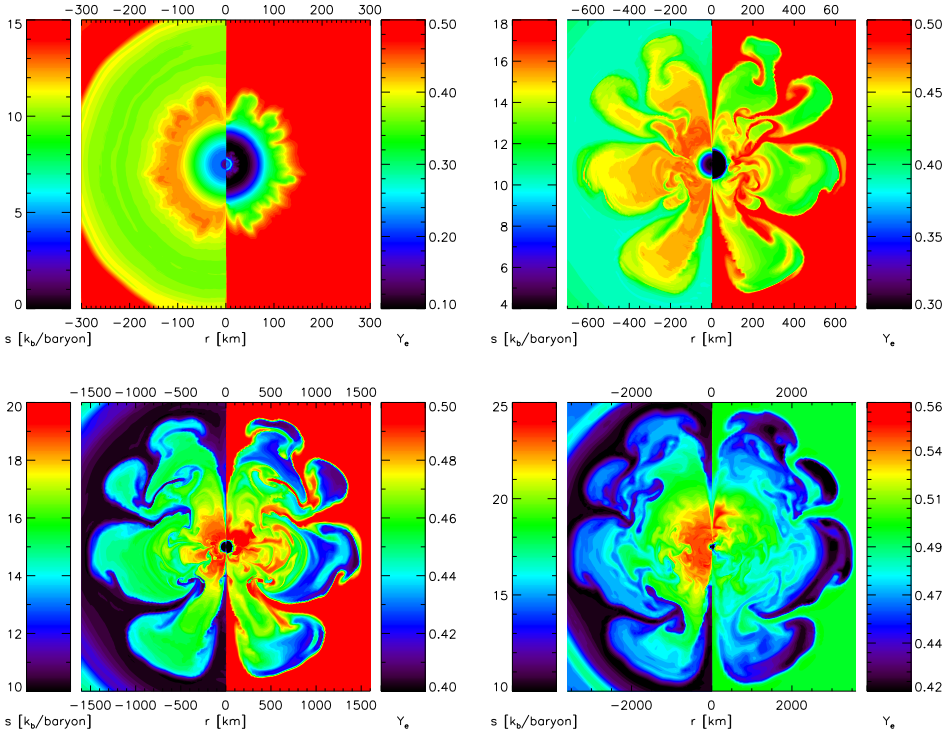


Figure 1. Snaphots showing the gas entropy (left half of panels) and the electron-to-nucleon ratio (right half of panels) for the explosion of a star with about nine solar masses. The plots correspond (from top left to bottom right) to times of 0.097, 0.144, 0.185, and 0.262 seconds after the formation of the supernova shock front and the onset of neutron star formation. Note the different radial scales of the four panels.

a critical issue here. The contributions of the NIC in Jülich to our resources were essential and allowed us in the past years to push forward into unexplored computational territory in modelling supernova explosions.

2 Computational Challenges

The modelling of supernova explosions is one of the greatest computational challenges in numerical astrophysics. Largely different time scales, varying between microseconds and seconds, and length scales, which extend from tens of meters to millions of kilometers, have to be resolved to follow neutrino interactions, nuclear burning, turbulent convection, and sound wave propagation in different regions of the collapsing core and of the ejected outer layers of an exploding star. This is not only computationally extremely demanding: half a second of evolution requires 500.000 time steps and in two spatial dimensions some 10^{18} floating point operations.

The problem is also hard to implement efficiently on massively parallel computers. In particular the neutrino transport module has resisted all such efforts so far and prevents us from the use of distributed memory architectures with many hundreds or thousands of

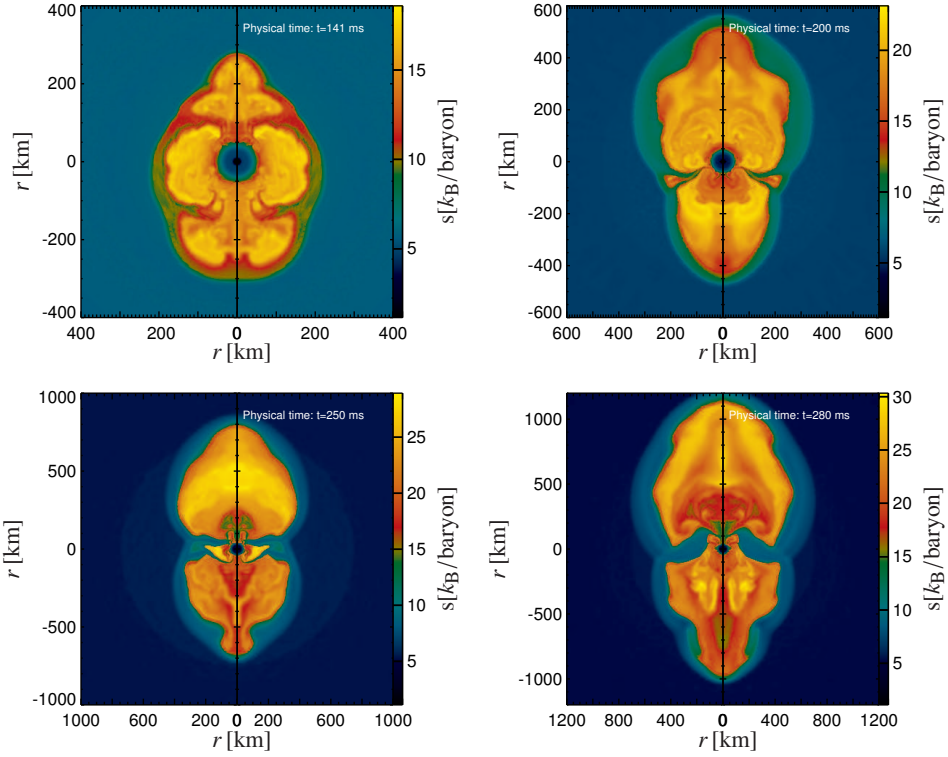


Figure 2. Snapshots of the gas entropy for the explosion of a star with about eleven solar masses at times 0.141, 0.200, 0.250 and 0.280 seconds after the launch of the supernova shock (top left to bottom right). The explosion develops a large asymmetry although the star is not rotating. Note the different radial scales of the four panels.

CPUs. Because the interaction time scales of neutrinos in neutron star matter are extremely short and the neutrino propagation happens with the speed of light after decoupling, the nonlinear transport equations of neutrinos, which as fermions are subject to phase-space blocking effects, need to be solved with fully implicit time stepping. In our current numerical implementation this leads to big, densly filled matrices that have to be inverted several times on every time level of the calculated evolution. This is computationally very expensive and defies easy parallelization. New algorithms, based on iterative multigrid solvers for hyperbolic systems of equations, are currently under development but are not yet available for full-scale supernova calculations. Their use, however, will be unavoidable in future three-dimensional models of supernovae.

The special needs of our current two-dimensional simulations, i.e., shared-memory nodes with powerful CPUs and continuous availability of these nodes for many months, cannot be fulfilled by many supercomputer centres. In particular the IBM p690 *Jump* of the NIC is therefore a highly valuable source of CPU time for our project.

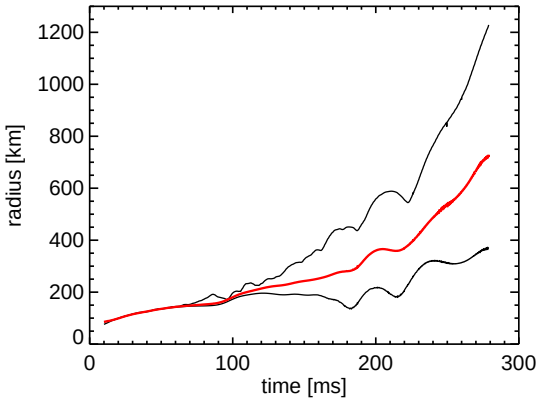


Figure 3. Maximum, average, and minimum radial position of the supernova shock front as functions of time for the explosion of a star with about eleven solar masses. Note the clear signature of several large-amplitude bipolar shock oscillations due to the standing accretion shock instability (SASI) before the blast takes off with an extreme 3:1 deformation.

3 Violent Gas Flows, Neutrino Heating, and Explosion

Our recent results for progenitor stars between 9 and 15 solar masses confirm the viability of the neutrino-heating mechanism for triggering supernova explosions. However, the inauguration of the blast happens not the way expected so far.

The behaviour of supernova progenitors with less than about 10 times the mass of the sun is clearly different from that of more massive stars. The former class of stars develops a core composed of oxygen, neon, and magnesium, not of iron, with an extremely steep density gradient at its surface. This allows the supernova shock front, which is launched at the moment when the neutron star begins to form in the collapsing stellar core, to expand continuously as it propagates into rapidly diluting infalling material. Behind the shock, neutrino heating deposits the energy that powers the explosion. Convective overturn develops in the neutrino heating layer behind the outgoing shock and imprints inhomogeneities on the ejecta, in entropy as well as composition (Fig. 1).

Evolved stars above roughly ten solar masses produce iron cores with a much more shallow density decline outside. Running into this denser material damps the initial expansion of the shock. Moreover, severe energy losses cause the shock to even stall. Different from previous calculations with simple grey neutrino diffusion, our more sophisticated models show that convection is also suppressed in the rapidly infalling matter behind the shock. It cannot become strong enough to give sufficient support for neutrino heating to cause the revival of the stagnant shock. Instead, another kind of nonradial hydrodynamic instability, the so-called standing accretion shock instability ("SASI"), obtains decisive influence on the shock evolution. With highest growth rates of the dipole and quadrupole modes, it leads to violent bipolar shock oscillations. This pushes the shock to larger radii and thus causes secondary convection and improves the conditions for efficient energy deposition by neutrinos, because the gas accreted through the stalled shock stays longer in the heating layer and is able to absorb more energy[5].

The presence of strong SASI oscillations is visible in both of our simulations for 11

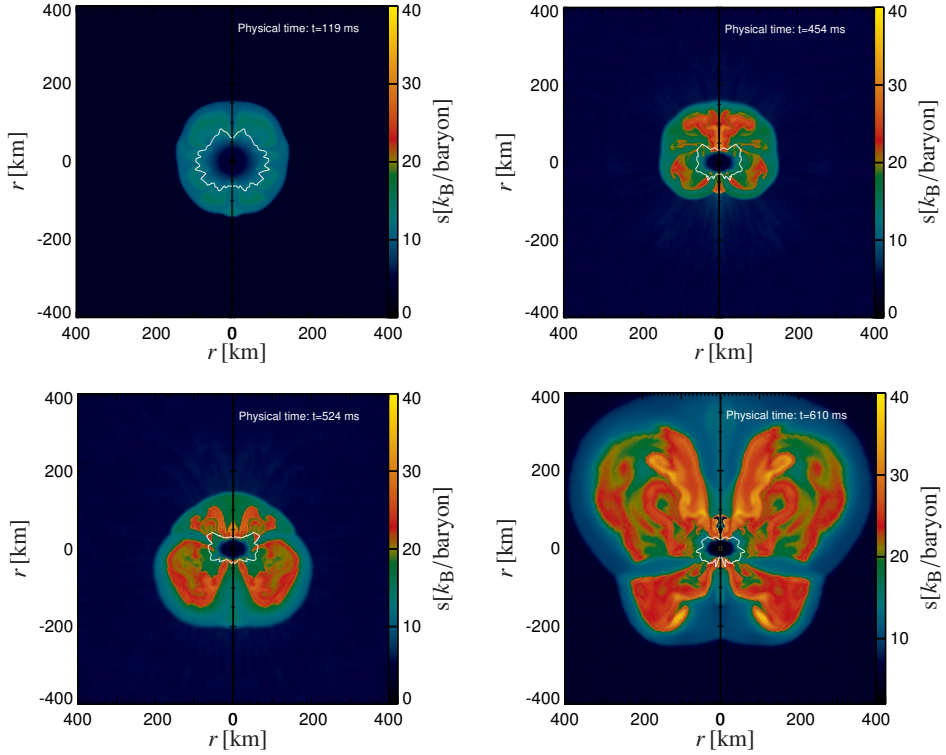


Figure 4. Snapshots of the gas entropy for the explosion of a star with 15 solar masses at times 0.119, 0.454, 0.524 and 0.610 seconds after the launch of the supernova shock (from top left to bottom right). A large north-south asymmetry establishes at the beginning of the explosion. The white line marks the lower boundary of the neutrino-heating region. Note the different radial scales of the four panels.

and 15 solar mass stars and turns out to be decisive for their ultimate explosion (Figs. 2–5). Different from the nine solar-mass star, where convection imposes a high-mode asymmetry pattern on the ejecta (see Fig. 1), the preferred growth of the dipole and quadrupole modes of the SASI leads to a large global anisotropy of the beginning supernova blast in the more massive progenitors[6,7]. The explosion may set in with a significant delay after the neutron star begins to form. For the most massive of the three investigated stars this happens 0.6 seconds later, which is much later than expected.

Our simulations have thus demonstrated the crucial role of the standing accretion shock instability in combination with neutrino heating for initiating supernova explosions. One of the consequences of this mechanism is a global asymmetry of the accelerating shock front and ejected gas. Scheck et al.[8] showed that such large anisotropies can leave the neutron star with a kick velocity that is in the range of the measured eigenvelocities of young pulsars. Moreover, the strongly deformed shock wave triggers mixing instabilities (Richtmyer-Meshkov and Rayleigh-Taylor) at the interfaces of the different composition shells of the exploding star after the passage of the outgoing shock wave. This can lead to large-scale mixing of the chemical elements between the deep interior and the outer stellar

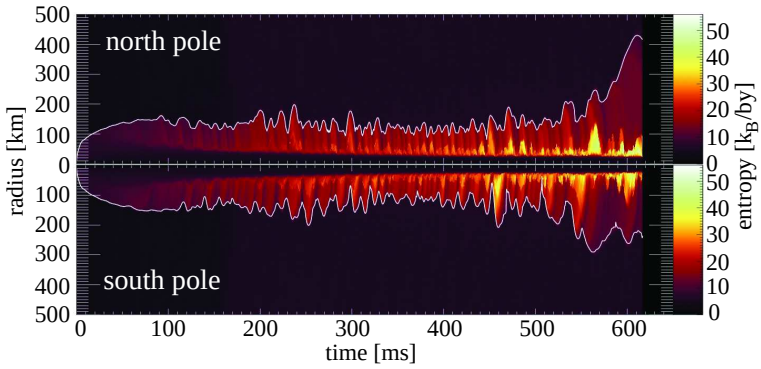


Figure 5. Shock positions near the north pole and near the south pole as functions of time for an exploding star with 15 solar masses. The gas entropy is colour coded. The plot shows the many cycles of quasi-periodic bipolar shock oscillations due to the standing accretion shock instability (SASI).

layers during the explosion, explaining self-consistently a variety of properties observed in well-monitored supernovae like the famous nearby Supernova 1987A[9].

4 Consequences and Future Perspectives

Establishing the progenitor-supernova connection and the final properties of the blast like its energy and the conditions for the formation of chemical elements (e.g., the proton-to-neutron ratio in the ejecta, see Fig. 1) requires many more simulations to be performed and each model to be conducted for many hundred milliseconds beyond the onset of the explosion. Following such a long evolution of highly dynamical behaviour is extremely demanding (also for the accuracy of the numerical scheme) and is one of reasons why these calculations pose a major computational challenge and require the significant amount of supercomputing resources we are also provided with by the NIC (see Section 2).

Recently, Burrows et al.[10,11] have found large-amplitude neutron star oscillations, so-called gravity waves, being excited by the violently turbulent gas flow around the forming compact remnant. In their simulations they saw a sizable amount of acoustic energy being radiated from the ringing neutron star, which thus acted as a transducer of gavitational binding energy released by accreted gas. In fact, this input of sonic energy was sufficient to even initiate the supernova. While this is an interesting alternative to trigger the explosion if neutrino heating fails, we do not observe such a dramatic effect. Yet, we clearly identify the presence of gravity modes in the neutron star. One of our goals for future long-time simulations is therefore a closer investigation of the question whether neutrino-driven explosions might receive an important contribution to their energy from the sound produced by the pulsations of the still anisotropically accreting nascent neutron star.

Acknowledgments

We are indebted to R. Buras, K. Kifonidis, E. Müller, and M. Rampp for fruitful collaborations. This project was supported by the Deutsche Forschungsgemeinschaft through the

Transregional Collaborative Research Centers SFB/TR 27 “Neutrinos and Beyond” and SFB/TR 7 “Gravitational Wave Astronomy”, the Collaborative Research Center SFB-375 “Astro-Particle Physics”, and the Cluster of Excellence “Origin and Structure of the Universe” (`http://www.universe-cluster.de`). The computations were only possible because of computer time on the IBM p690 of the John von Neumann Institute for Computing (NIC) in Jülich, on the national supercomputer NEC SX-8 at the High Performance Computing Center Stuttgart (HLRS) under grant number SuperN/12758, on the IBM p690 of the Computer Center Garching (RZG), on the sgi Altix 4700 of the Leibniz-Rechenzentrum (LRZ) in Munich, and on the sgi Altix 3700 of the MPI for Astrophysics. We also acknowledge support by AstroGrid-D, a project funded by the German Federal Ministry of Education and Research (BMBF) as part of the D-Grid initiative.

References

1. S.A. Colgate and R.H. White, *The Hydrodynamic Behavior of Supernovae Explosions*, Astrophys. Journal **143**, 626–681, 1966.
2. H.A. Bethe and J.R. Wilson, *Revival of a Stalled Supernova Shock by Neutrino Heating*, Astrophys. Journal **295**, 14–23, 1985.
3. H.-Th. Janka, K. Langanke, A, Marek, G. Martínez-Pinedo, and B. Müller, *Theory of Core-Collapse Supernovae*, Phys. Reports **442**, 38–74, 2007.
4. M. Rampp and H.-Th. Janka, *Radiation Hydrodynamics with Neutrinos: Variable Eddington Factor Method for Core-Collapse Supernova Simulations*, Astron. Astrophys. **396**, 361–392, 2002.
5. L. Scheck, H.-Th. Janka, T. Foglizzo, and K. Kifonidis, *Multidimensional Supernova Simulations with Approximative Neutrino Transport. II. The Advective-Acoustic Cycle in the Supernova Core*, Astron. Astrophys., **in press**, 2007, arXiv:0704.3001.
6. R. Buras, H.-Th. Janka, M. Rampp, and K. Kifonidis, *Two-Dimensional Hydrodynamic Core-Collapse Supernova Simulations with Spectral Neutrino Transport. II. Models for Different Progenitor Stars*, Astron. Astrophys. **457**, 281–308, 2006.
7. A. Marek and H.-Th. Janka, *Delayed Neutrino-Driven Supernova Explosions Aided by the Standing Accretion-Shock Instability*, Astrophys. Journal, **submitted**, 2007, arXiv:0708.3372.
8. L. Scheck, K. Kifonidis, H.-Th. Janka, and E. Müller, *Multidimensional Supernova Simulations with Approximative Neutrino Transport. I. Neutron Star Kicks and the Anisotropy of Neutrino-Driven Explosions in Two Spatial Dimensions*, Astron. Astrophys. **457**, 963–986, 2006.
9. K. Kifonidis, T. Plewa, L. Scheck, H.-Th. Janka, and E. Müller, *Non-Spherical Core Collapse Supernovae. II. The Late-Time Evolution of Globally Anisotropic Neutrino-Driven Explosions and Their Implications for SN 1987A*, Astron. Astrophys. **453**, 661–678, 2006.
10. A. Burrows, E. Livne, L Dessart, C.D. Ott, and J. Murphy, *A New Mechanism for Core-Collapse Supernova Explosions*, Astrophys. Journal **640**, 878–890, 2006.
11. A. Burrows, E. Livne, L Dessart, C.D. Ott, and J. Murphy, *Features of the Acoustic Mechanism of Core-Collapse Supernova Explosions*, Astrophys. Journal **655**, 416–433, 2007.

Formation and Evolution of Black Holes in Galactic Nuclei and Star Clusters

Rainer Spurzem[1], Peter Berczik[1], Ingo Berentzen[1], David Merritt[2], Miguel Preto[1], Pau-Amaro Seoane[3], Andrea Borch[1], Jonathan M.B. Downing[1], Christoph Eichhorn[1,4], Andreas Ernst[1], José Fiestas[1], Andreas Just[1], Chingis Omarov[1,5], and Oliver Porth[1]

[1] Astronomisches Rechen-Institut, Zentrum für Astronomie
Univ. Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany
E-mail: {spurzem, berczik, iberent, borch, downin, aernst, fiestas, miguelp, oporth}@ari.uni-heidelberg.de

[2] Dept. of Physics, Rochester Institute of Technology
85 Lomb Memorial Drive, Rochester, NY 14623-5604, USA
E-mail: merritt@astro.rit.edu

[3] Max Planck Institute for Gravitational Physics (Albert Einstein Institute)
Am Mühlenberg 1, 14476 Golm, Germany
E-mail: pau@aei.mpg.de

[4] Institut für Raumfahrtsysteme, Univ. Stuttgart, Pfaffenwaldring 31, 70550 Stuttgart, Germany
E-mail: eichhorn@irs.uni-stuttgart.de

[5] Fessenkov Astrophysical Institute, 050020 Almaty, Kazakhstan
E-mail: {omarov@aphi.kz, chingis.omarov@gmail.com}

This project studies the formation, growth, and co-evolution of single and multiple supermassive black holes (SMBHs) and compact objects like neutron stars, white dwarfs, and stellar mass black holes in galactic nuclei and star clusters, focusing on the role of stellar dynamics. In this paper we focus on one exemplary topic out of a wider range of work done, the study of orbital parameters of binary black holes in galactic nuclei (binding energy, eccentricity, relativistic coalescence) as a function of initial parameters. In some cases the classical evolution of black hole binaries in dense stellar systems drives them to surprisingly high eccentricities, which is very exciting for the emission of gravitational waves and relativistic orbit shrinkage. Such results are interesting to the emerging field of gravitational wave astronomy, in relation to a number of ground and space based instruments designed to measure gravitational waves from astrophysical sources (VIRGO, Geo600, LIGO, LISA). Our models self-consistently cover the entire range from Newtonian dynamics to the relativistic coalescence of SMBH binaries.

1 Introduction

MBH formation and their interactions with their host galactic nuclei is an important ingredient for our understanding of galaxy formation and evolution in a cosmological context, e.g. for predictions of cosmic star formation histories or of MBH demographics (to predict events which emit gravitational waves). If galaxies merge in the course of their evolution, there should be either many binary or even multiple black holes, or we have to find out what happens to black hole multiples in galactic nuclei, e.g. whether they come close enough together to merge under emission of gravitational waves, or whether they eject each other in gravitational slingshot. For numerical simulations of the problem all models

depend on an unknown scaling behaviour, because the simulated particle number is not yet realistic due to limited computing power[25, 26, 22, 6]. Dynamical modelling of non-spherical dense stellar systems (with and without central BH) is even less developed than in the spherical case. Here we present a set of numerical models of the formation and evolution of binary black holes in rotating galactic nuclei. Since we are interested in the dynamical evolution of MBH binaries in their final phases of evolution (the last parsec problem) we somehow abstract from the foregoing complex dynamics of galactic mergers. We assume that after some violent dynamic relaxation a typical initial situation consists of a spherical or axisymmetric coherent stellar system (galactic nucleus), where fluctuations in density and potential due to the galaxy merger have decayed, which is reasonable on an (astrophysically) short time scale of a few ten million years. The MBHs, which were situated in the centre of each of the previously merged galaxies, are located at the boundary of the dense stellar core, some few hundred parsec apart. This situation is well observable[14].

According to the standard theory, the subsequent evolution of the black holes is divided in three intergradient stages[5]: 1. Dynamical friction causes a transfer of the black holes' kinetic energy to the surrounding field stars, the black holes spiral to the centre where they form a binary. 2. While hardening, the effect of dynamical friction reduces and the evolution is dominated by superelastic scattering processes, that is the interaction with field stars closely encountering or intersecting the binaries' orbit, thereby increasing the binding energy. 3. Finally the black holes coalesce throughout the emission of gravitational radiation, potentially detectable by the planned space-based gravitational wave antennae LISA.

In this paper, the behaviour of the orbital elements of a black hole binary in a dense stellar system is investigated in a self-consistent way from the beginning till the relativistic merger and its emission of gravitational waves. The evolution of the eccentricity has been discussed for some time[21, 12, 25, 6, 22]. According to Peters & Mathews and Peters[28, 27] the timescale of coalescence due to the emission of gravitational radiation is given by

$$t_{gr} = \frac{5}{64} \frac{c^5 a_{gr}^4}{G^3 M_1 M_2 (M_1 + M_2) F(e)} \tag{1}$$

wherein M_1, M_2 denote the black hole masses, a_{gr} the characteristic separation for gravitational wave emission, G the gravitational constant, c the speed of light and

$$F(e) = (1 - e^2)^{-7/2} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right) \tag{2}$$

a function with strong dependence on the eccentricity e. Thus the coalescence time can shrink by several orders of magnitude if the eccentricity is high enough, resulting in a strengthened burst of gravitational radiation. Highly eccentric black hole binaries would represent appropriate candidates for forthcoming verification of gravitational radiation through the planned mission of the Laser Interferometer Space Antenna mission LISA.

2 Numerical Method, Initial Models

The simulations have been performed using NBODY6++, a parallelized version of Aarseth's NBODY6[1, 35, 2]. The code includes a Hermite integration scheme, KS-regularization[15] and the Ahmad-Cohen neighbour scheme[4]. No softening of the interaction potential of any two bodies is introduced; this allows an accurate treatment of the

effects due to superelastic scattering events, which play a crucial part in black hole binary evolution and require a precise calculation of the trajectories throughout the interaction. The code and its parallel performance has been described in detail in this series and elsewhere[35, 13]. The survey has been carried out for a total particle number of up to $N = 1\ 000\ 000$ including two massive black holes with $M_1 = M_2 = 0.01$ embedded in a dense stellar system of equal-mass particles $m_* \approx 1.0 \cdot 10^{-6}$. The total mass of the system is normalized to unity. The initial stellar distribution was taken from generalized King models with and without rotation[16, 19, 11].

3 Simulations

3.1 Newtonian Evolution of the Binary Black Hole

In the first evolutionary stage, each black hole individually suffers dynamical friction with the surrounding low mass stars, which is the main process of losing energy. The role of dynamical friction decreases when a permanently bound state occurs, as the dynamical friction force acts preliminary on the motion of the now formed binary rather than on the individual black holes. Superelastic scattering events of field stars at the binary then become more and more important for the reduction of its energy. The process sustains an ongoing "hardening" of the binary (shrinking of semi-major axis and increase of energy) and also in many cases a high eccentricity. While the hardening rates are well understood[33, 30] and do not depend much on the initial parameters of the preceding galactic merger, this is not as clear for the eccentricity, which depends on initial conditions at least to some extent[10].

In a spherically symmetric system the binary hardening would stall after a few crossing times, because loss-cone orbits of stars, which come close to the cenral SMBH binary will be depleted, and replenishment takes place only on a much longer relaxation time. This effect is more dramatic for systems with large particle number, because the relaxation time increases strongly, and is depicted on the top panel of Fig. 1; it has been claimed that in real galaxies with their very large particle numbers therefore the SMBH binary will not reach relativistic coalescence. This situation was relaxed from two sides, first by a careful analysis of loss-cone refilling time scales combining direct N-body and Fokker-Planck models[23], and by looking for a moderately rotating, axisymmetric galactic nucleus[7], where the loss cone remains full even for large particle numbers (see lower panel in Fig. 1). Since some degree of perturbation of a spherical model is quite natural for a remnant after galactic mergers, many of them might even be triaxial rather than axisymmetric, the stalling problem does not exist anymore.

4 Relativistic Dynamics of Black Holes in Galactic Nuclei

4.1 Introduction

Relativistic stellar dynamics is of paramount importance for the study of a number of subjects. For instance if we want to have a better understanding of what the constraints on alternatives to supermassive black holes are; in order to canvass the possibility of ruling out stellar clusters, one must do detailed analysis of the dynamics of relativistic clusters. Furthermore the dynamics of compact objects around SMBH and of multiple SMBH in

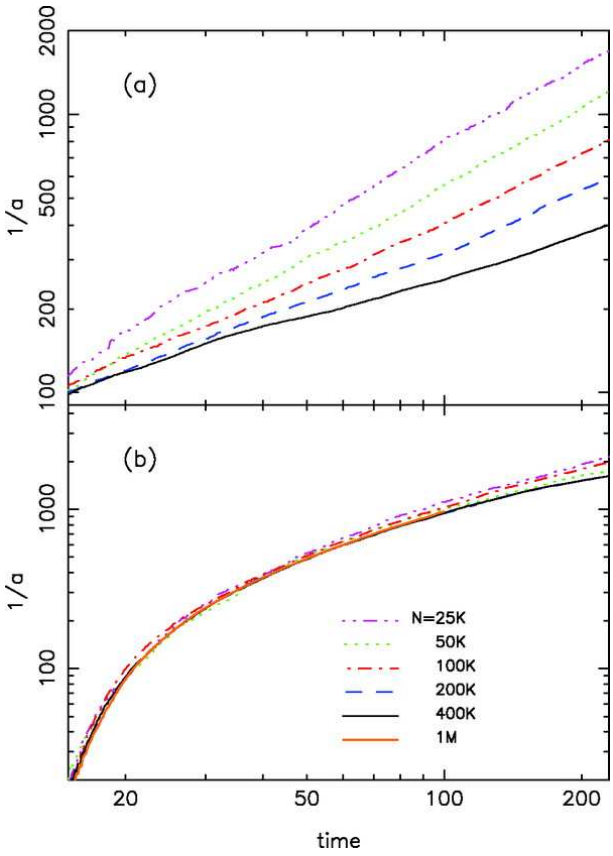


Figure 1. Evolution of the inverse semi-major axis of a black hole binary in direct N-body models with varying particle number N - top panel: for spherically symmetric systems a stalling occurs; - bottom panel: for axisymmetric rotating systems there is no sign of stalling. Compare[7].

galactic nuclei requires the inclusion of relativistic effects. Our current work deals with the evolution of two SMBHs, bound to each other, and looking at the phase when they get close enough to each other that relativistic corrections to Newtonian dynamics become important, which ultimately lead to gravitational radiation losses and coalescence.

Efforts to understand the dynamical evolution of a stellar cluster in which relativistic effects may be important have been already done by[17],[32],[31] and[18]. In the earlier work $1\mathcal{PN}$ and $2\mathcal{PN}$ terms were neglected[18] and the orbit-averaged formalism[27] used. We describe here a method to deal with deviations from Newtonian dynamics more rigorously than in most existing literature (but compare[24,3], which are on the same level of PNaccuracy). We modified the NBODY6++ code to allow for post-Newtonian ($\mathcal{PN}$) effects of two particles getting very close to each other, implementing in it the $1\mathcal{PN}$, $2\mathcal{PN}$ and $2.5\mathcal{PN}$ corrections fully from[34,37].

4.2 Method: Direct Summation NBODY with Post-Newtonian Corrections

The version of direct summation NBODY method we employed for the calculations, NBODY6++, includes the *KS regularisation*. This means that when two particles are tightly bound to each other or the separation among them becomes smaller during a hyperbolic encounter, the couple becomes a candidate for a regularisation in order to avoid problematical small individual time steps[15]. We modified this scheme to allow for relativistic corrections to the Newtonian forces by expanding the acceleration in a series of powers of $1/c$ in the following way[9,34]:

$$\underline{a} = \underbrace{\underline{a}_0}_{\text{Newt.}} + \underbrace{\underbrace{c^{-2}\underline{a}_2}_{1\mathcal{PN}} + \underbrace{c^{-4}\underline{a}_4}_{2\mathcal{PN}}}_{\text{periastron shift}} + \underbrace{\underbrace{c^{-5}\underline{a}_5}_{2.5\mathcal{PN}}}_{\text{grav. rad.}} + \mathcal{O}(c^{-6}), \tag{3}$$

where $\underline{a}$ is the acceleration of particle 1, $\underline{a}_0 = -Gm_2\underline{n}/r^2$ is the Newtonian acceleration, G is the gravitation constant, m_1 and m_2 are the masses of the two particles, r is the distance of the particles, $\underline{n}$ is the unit vector pointing from particle 2 to particle 1, and the $1\mathcal{PN},\ 2\mathcal{PN}$ and $2.5\mathcal{PN}$ are post-Newtonian corrections to the Newtonian acceleration, responsible for the pericenter shift ($1\mathcal{PN},\ 2\mathcal{PN}$) and the quadrupole gravitational radiation ($2.5\mathcal{PN}$), correspondingly, as shown in Eq. (3). As an example we give the expressions for the $1\mathcal{PN}$ and $2.5\mathcal{PN}$ terms, for $2\mathcal{PN}$ see the cited literature[34]:

$$\underline{a}_2 = \frac{Gm_2}{r^2}\Big\{\underline{n}\Big[-v_1^2 - 2v_2^2 + 4v_1v_2 + \tfrac{3}{2}(nv_2)^2 + 5\big(\tfrac{Gm_1}{r}\big) + 4\big(\tfrac{Gm_2}{r}\big)\Big] + (\underline{v}_1 - \underline{v}_2)\big[4nv_1 - 3nv_2\big]\Big\} \tag{4}$$

$$\underline{a}_5 = \tfrac{4}{5}\tfrac{G^2m_1m_2}{r^3}\Big\{(\underline{v}_1 - \underline{v}_2)\Big[-(\underline{v}_1 - \underline{v}_2)^2 + 2\big(\tfrac{Gm_1}{r}\big) - 8\big(\tfrac{Gm_2}{r}\big)\Big] + \underline{n}(nv_1 - nv_2)\big[3(\underline{v}_1 - \underline{v}_2)^2 - 6\big(\tfrac{Gm_1}{r}\big) + \tfrac{52}{3}\big(\tfrac{Gm_2}{r}\big)\big]\Big\}. \tag{5}$$

In the last expressions $\underline{v}_1$ and $\underline{v}_2$ are the velocities of the particles. For simplification, we have denoted the vector product of two vectors, $\underline{x}_1$ and $\underline{x}_2$, like x_1x_2. We integrated our correcting terms as external forces into the *two-body KS regularisation* scheme which requires to compute their time derivatives in the Hermite scheme as well (not shown in equations here for brevity).

4.3 First Results

In Fig. 2 the impact of relativistic, Post-Newtonian dynamics to the separation of the binary black holes in our simulations is seen. The curve deviates from the Newtonian results when gravitational radiation losses set in and cause a sudden coalescence ($1/a \to \infty$) at a finite time. Gravitational radiation losses are supported by the high eccentricity of the SMBH binary. It is interesting to note that the inclusion or exclusion of the conservative 1 and 2PN terms changes the coalescence time considerably. Details of these results will be published elsewhere[8].

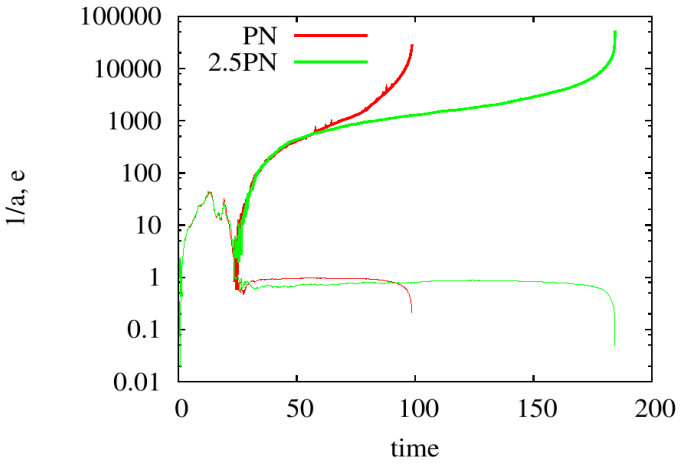


Figure 2. Effect of Post-Newtonian (PN) relativistic corrections on the dynamics of black hole binaries in galactic nuclei, plotted are inverse semi-major axis and eccentricity as a function of time. The red line uses the full set of PN corrections, while the green line has been obtained by artificially only using the dissipative $\mathcal{PN}$ 2.5 terms. Here $c = 457$ has been chosen in model units, see more details in[8].

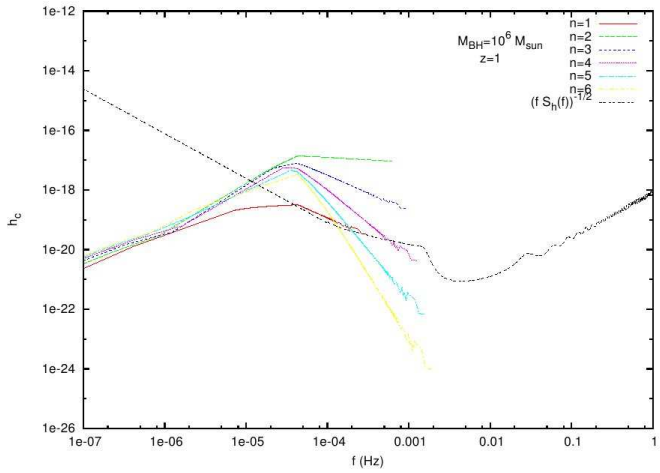


Figure 3. Locus of SMBH binaries from our simulations in the LISA sensitivity diagram during their final inspiral and coalescence. Plotted is dimensionless strain versus frequency, for all relevant harmonics (circular orbit: $n = 2$ dominates), and the LISA sensitivity curve. We have selected as an example here the signal expected from a SMBH binary of one million solar masses located at a redshift of $z \approx 1$. Further details will be published elsewhere[29].

Once the SMBH binary starts to dramatically lose binding energy due to gravitational radiation its orbital period will drop from a few thousand years to less than a year very quickly (timescale much shorter than the dynamical time scale in the galactic center, which

defines our time units). Then the SMBH binary will enter the LISA band, i.e. its gravitational radiation will be detectable by LISA. LISA, Laser Interferometer Space Antenna, is a system of three space probes with laser interferometers to measure gravitational waves, see e.g. `http://lisa.esa.int/`. Once our SMBH binary decouples from the rest of the system we just follow its relativistic two-body evolution, starting with exactly the orbital parameters (eccentricity!) as they came out from the N-body model. It is then possible to predict the gravitational radiation of our SMBH binary relative to the LISA sensitivity curve, which is depicted in Fig. 3. Plotted are different harmonics of the gravitational radiation, for the circular orbit $n = 2$ is dominant, while for eccentric orbits higher harmonics are stronger[28,27]. One can see in the plot how the SMBH binary enters into the LISA sensitivity regime with some eccentricity.

5 Summary

We have shown that supermassive black hole binaries in galactic nuclei may reach the stalling barrier and will reach the relativistic coalescence phase in a timescale shorter than the age of the universe. A gravitational wave signal expected for the LISA satellite from these SMBH binaries is expected, in particular due to the high eccentricity of the SMBH binary when entering the relativistic coalescence phase. Our models cover self-consistently the transition from the Newtonian dynamics to the situation when relativistic, Post-Newtonian ($\mathcal{PN}$) corrections start to influence the relative SMBH motion. After the shrinking time scale became very short the binary decouples from the rest of the galactic nucleus and can be treated as a relativistic two-body problem. We follow this evolution formally to the coalescence of the two black holes using $\mathcal{PN}$ terms of up to order $2.5\mathcal{PN}$ and determine the gravitational wave emission in different modes relative to the LISA sensitivity curve.

This paper has only selected one of the highlights of a number of applications of our parallel direct N-body code NBODY6++, because of constraints of space. Other projects followed here are the detailed modelling of populations and spectra, as well as gravitational wave emission by neutron stars, black holes and white dwarfs from globular clusters (for ground based detectors such as VIRGO, LIGO, GEO600) (by A. Borch, J. Downing), the study of inspiralling globular clusters onto the SMBH in our Galaxy (by A. Ernst, A. Just[36]), star-disk interactions in thick accretion disks around SMBH (by C. Omarov) and modelling of loss cones and tidal disruption near a single black hole (by O. Porth). Last but not least direct N-body models of galactic nuclei need to be complemented still by statistical models to reach realistically high particle numbers - here we use an orbit averaged 2D Fokker-Planck equation (by J. Fiestas).

Acknowledgments

Computing time at NIC Jülich on the IBM Jump is acknowledged. Financial support comes partly from Volkswagenstiftung, German Science Foundation (DFG) via SFB439 at the University of Heidelberg and Schwerpunktprogramm 1177 (Project ID Sp 345/17-1) 'Black Holes Witnesses of Cosmic History' It is a pleasure to acknowledge many enlightening discussions and support by Sverre Aarseth, very useful interactions about relativistic dynamics with A. Gopakumar and G. Schäfer. Part of the work presented here overlaps with the contents of the DEISA project 'Einstein'.

References

1. Aarseth S.J., 1999a, Publ. Astr. Soc. Pac. 111, 1333.
2. Aarseth, S., 2003, *Gravitational N-body simulations* (Cambridge University Press, Cambridge), p.173.
3. Aarseth, S. J., Monthly Notices of the Royal Astronomical Society **378**, 285 (2007).
4. Ahmad A., Cohen L., Journal of Computational Physics 1973, 12, 349.
5. Begelman M.C., Blandford R.D., Rees M.J., 1980, Nature, 287, 307.
6. Berczik, P., Merritt, D., Spurzem, R. 2005, ApJ 633, 680.
7. Berczik, P., Merritt, D., Spurzem, R., Bischof, H.-P., 2006, ApJ 642, L21.
8. Berentzen I., Berczik P., Merritt, D., Spurzem R., Preto M., 2008, in prep..
9. T. Damour & N. Dereulle, Phys. Lett.,1987, 87, 81.
10. Eichhorn C., Amaro-Seoane P., Spurzem R., 2008, in prep..
11. Einsel C., Spurzem R., 1999, 1999, MNRAS, 302, 81.
12. Hemsendorf M., Sigurdsson S., Spurzem R., 2002, ApJ, 581, 1256.
13. Khalisi, E., Omarov, C.T., Spurzem, R., Giersz, M., Lin, D.N.C. 2003, in "High Performance Computing in Science and Engineering", Springer Vlg., pp. 71-89.
14. Komossa S., Burwitz V., Hasinger G., Predehl P., Kaastra J.S., Ikebe, Y., 2003, ApJ, 582, L15.
15. Kustaanheimo P., Stiefel E., Journal für die reine und angewandte Mathematik 1965, 218, 204.
16. Lagoute C., Longaretti P.-Y., 1996, A&A, 308, 441.
17. H. M. Lee, ApJ, 319, 801, 1987.
18. Lee, M. H., Astrophysical Journal **418**, 147 (1993) .
19. Longaretti P.-Y., Lagoute C., 1996, A&A, 308, 453.
20. J. Makino & S. J. Aarseth, Publ. Astron. Soc. Japan, 44, 141, 1992.
21. Makino J., Fukushige T., Okumura S.K., Ebisuzaki T., 1993, PASJ 45, 303.
22. Makino, J. & Funato, Y. 2004, ApJ, 602, 93.
23. Merritt, D., Mikkola, S., Szell, A., ArXiv e-prints **705**, arXiv:0705.2745 (2007) .
24. Mikkola, S., Merritt, D., ArXiv e-prints **709**, arXiv:0709.3367 (2007).
25. Milosavljević M., Merritt D., 2001, ApJ, 563, 34.
26. Milosavljević, M. & Merritt, D. 2003, ApJ, 596, 860.
27. Peters P.C., 1964, Phys.Rev., 136, B1224.
28. Peters, P.C. & Mathews, J. 1963, Phys. Rev. 131, 435.
29. Preto M., et al. 2008, in prep..
30. Quinlan, G. D., New Astronomy **1**, 35 (1996).
31. Quinlan, G. D., Shapiro, S. L., Astrophysical Journal **356**, 483 (1990).
32. Quinlan, G. D., Shapiro, S. L., Astrophysical Journal **343**, 725 (1989).
33. Sesana, A., Haardt, F., Madau, P., Astrophysical Journal **660**, 546 (2007).
34. M. H. Soffel, 1989, "Relativity in Astrometry, Celestial Mechanics and Geodesy", Springer-Verlag Berlin, Heidelberg, New York.
35. Spurzem R., 1999, Journal of Computational and Applied Mathematics, 109, 407.
36. Ernst A., Glaschke P., Fiestas J., Just A., Spurzem R., 2007, MNRAS 377, 465.
37. Kupi G., Amaro-Seoane P., Spurzem R., 2006, MNRAS 371, L45.

Biophysics

Harold A. Scheraga

Baker Laboratory of Chemistry and Chemical Biology, Cornell University
Ithaca, NY 14853-1301, USA
E-mail: has5@cornell.edu

Computation in the life sciences is currently a frontier field, relying on large-scale computing to treat the properties of bio-macromolecules. Computation helps in two different, but related ways: (1) to rationalize observed experimental results in terms of intermolecular interactions, and (2) to obtain information about equilibrium and dynamic properties of systems that are not yet amenable to experimentation. The biological systems treated in this NIC symposium consist of proteins and an even larger system, namely cells.

Because of the complexity of these systems, it is not yet possible, even with modern supercomputers, to treat them at the all-atom level, in which the time evolution of the coordinates and velocities of every atom in the system is describable. Therefore, coarse-grained or mesoscopic models must be devised to render the computations feasible. However, much ingenuity must go into the formulation of such models, and the algorithms to treat them, to ensure that they capture the essential features of the all-atom systems from which they are derived.

The three papers presented here are elegant examples in each of which a very important biological problem is treated at the coarse-grained level with the help of the computing facilities at the John von Neumann Institute for Computing. It is clear that, without such computing facilities, the progress described in these papers could not have been achieved.

The paper of Liwo et al describes the application of a physics-based united-residue force field (UNRES) to simulate the structure and long-time dynamics as a newly-synthesized polypeptide chain of amino acids folds into the unique three-dimensional structure of a biologically active protein. The required research on this problem involves the formulation of a realistic potential-energy function (UNRES) and its use in a molecular dynamic treatment of protein folding.

Another application to a protein is the paper of Geibig et al that considers the metal-binding properties of an enzyme, haloperoxidase. This paper deals with how the enzyme discriminates among different metals and influences the catalytic activity of the metal-protein complex. By DFT studies on several model systems of haloperoxidases, this paper explores the mechanism of the action of these enzymes, and postulates an explanation for the discrimination among different metals by the enzyme.

The paper of Elgeti and Gompper treats an even larger biological system, namely, a sperm cell. Here, the interest is in the mobility of sperm cells in carrying out their biological function. The dynamics of such a system is studied with a mesoscale hydrodynamic simulation technique to treat multi-particle collision dynamics, with the sperm tail modelled by a crane-like structure. Details of the trajectory are thereby determined by considering the structural details of the sperm cell.

In summary, these three papers represent advances in frontier biological problems, treated by physical theory and computations which are facilitated by access to modern supercomputers.

Hydrodynamics of Active Mesoscopic Systems

Jens Elgeti and Gerhard Gompper

Institut für Festkörperforschung
Research Centre Jülich, 52425 Jülich, Germany
E-mail: {j.elgeti, g.gompper}@fz-juelich.de

The dynamics of active mesoscopic systems is studied by multi-particle collision dynamics, a mesoscale hydrodynamic simulation technique. Sperm motion serves as a model system par excellence. The sperm tail is modelled by a crane-like structure. It is shown that helical trajectories are generated by a chiral structure of the sperm. The connection between sperm geometry and parameters of the trajectory is determined.

1 Introduction

Many eucaryotic cells and bacteria move and navigate actively in their environment. Some well-known examples are E.coli, which swims by a rotating helical filament in search for food, fibroblasts, which crawl to close a wound, and sperm cells, which swim with a beating flagellum to find the egg. The catalog of cell motion is long, each form especially suited for the given task of the cell.

Cell crawling, where cells adhere to surfaces and extend filopodia to pull themselves forward, is a slow process, with typical velocities around 10 $\mu m/min$. We are interested here in the faster forms of mesoscopic swimming. Sperm cells are the model systems par excellence, with swimming velocities around 50 $\mu m/s$. Swimming on mesoscopic length scales of nano- to micro-meters, is a lot different than on the macroscopic length scales familiar to us[1]. Hydrodynamics is in general described by the Navier Stokes equation. The Reynolds number $Re = \rho u l/\eta$, where ρ is the density and η the viscosity of the fluid, and u and l are characteristic length and velocity scales, determines the importance of the non-linear inertia terms. In the low Reynolds number regime, viscous forces dominate over inertia, leading to the linear and time-reversible Stokes equation. The consequences of a low Reynolds number are quite surprising. For example, a coloured drop can be deposited in a high-viscosity fluid between two concentric cylinders. When the inner cylinder is turned several times to mix the fluids, and susequently the same number of times in the opposite direction, the result is *not* a homogeneous mixture – instead the drop is reformed![2]. This effect renders time-reversible swimming mechanisms impossible. A nice review of *swimming at low Reynolds numbers* can be found in Ref.[1].

Sperm cells propel themselves forward by a snake-like motion of their tail, the flagellum, towards the egg. The best-known example is clearly human sperm, but almost all higher species, from the sea urchin to the elephant, produce sperm. These sperm all look remarkably alike. The flagellum may have a different length, but its underlying cytoscelatal structure is well preserved throughout animal live. The direction of sperm motion is controlled by chemotaxis, and has been studied extensively. Recently, the biochemical signaling cascade of chemotaxis of Sea Urchin spermatozoa has been unraveled[3]. Our work is focused on the *hydrodynamic* effects in these systems.

Cilia are hair-like extensions from the cell that propel fluid by a whip-like motion. Cilia are even more abundant in nature than sperm. The Paramecium is covered with thousands of cilia to propel it through the fluid, the mammalian lung is equipped with cilia to propel mucus out of the system, and in the female reproductive tract cilia move the egg.

Typically cilia appear in large arrays, so that hydrodynamic interactions between multiple active compounds play an essential role. It has been proposed, but not yet proven, that hydrodynamic interactions lead to metachronal waves, a self organized pattern found on ciliated surfaces. In a large array, cilia do not beat synchronously or randomly, but in a well defined wave-like pattern.

Activity on mesoscopic length scales is what is uniting these systems. They consist of intrinsically active, micrometer-sized filaments in fluid suspension. Activity is not imposed by outside fields, but generated within the objects themselves.

2 Mesoscale Hydrodynamics

Hydrodynamics plays an essential role in mesoscopic systems. Clearly the systems are too large (several μm) to model the water explicitly. Furthermore the hydrodynamics has only to be resolved on the relevant length scales of the system. This calls for a coarse-grained description of the fluid.

Several mesoscale simulation techniques to describe the hydrodynamics of complex fluids have been proposed in recent years, such as Lattice-Boltzmann, Dissipative Particle Dynamics and Multi-Particle-Collision Dynamics. The basic idea in all these approaches is to employ a highly simplified dynamics on the microscale, but to respect the relevant conservation laws for mass, momentum and energy, such that hydrodynamical behaviour emerges naturally on larger length scales.

Multi-Particle Collision Dynamics (MPC) is one such technique[4–6]. In MPC, the fluid is represented by N point particles in continuous space. The particles travel with continuous velocities, and interact with a coarse-grained multi-particle interaction. The dynamics evolves in two steps. In the *streaming step*, the particles move ballistically for a collision-time h,

$$\vec{r}_i(t+h) = \vec{r}_i(t) + h\vec{v}_i(t). \tag{1}$$

In the *collision step*, the particles are sorted into the cells of a cubic lattice with lattice constant a and their velocities, relative to the centre-of-mass velocity of all particles in the cell, are rotated around a random axis by an angle α,

$$\vec{v}_i(t+h) = \vec{v}_{cm}(t) + \mathfrak{R}(\vec{v}_i(t) - \vec{v}_{cm}(t)), \tag{2}$$

where $\mathfrak{R}$ is a rotation matrix. The direction of rotation is chosen independently for each box at each time step, while the rotation angle is typically constant and used as a parameter to tune viscosity and Schmidt number $S_c = \eta/\rho D$, the ratio of momentum to mass diffusivity (D is the diffusion constant). It has been shown that large collision angles and small time steps are best suited for simulating fluid-like behaviour at low Reynolds numbers[7,8].

To ensure Galilean invariance, the collision grid has to be shifted between subsequent collision steps[6,9].

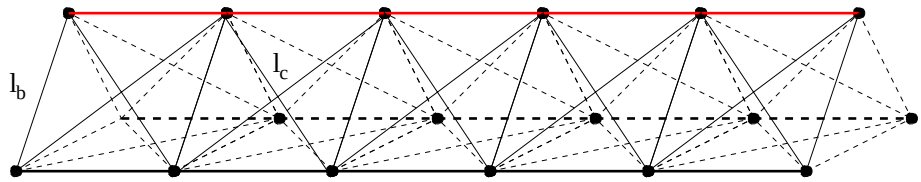


Figure 1. The axoneme is modelled as three polymer rods, interconnected by harmonic springs. Bond length $l_b = a/2$, $l_c = 1/\sqrt{2}$.

Both, the collision step and the streaming step conserve momentum and energy explicitly, so that hydrodynamic behaviour emerges automatically on large length scales. Due to the particle-based description of the fluid, coupling to other embedded mesoscale objects is straightforward. In particular Molecular Dynamics is well suited to simulate immersed objects like polymers or vesicles.

One big advantage of this method is its computational efficiency, due to the simple interactions between the particles. Furthermore, the largest part of the iterations are independent of each other and the interactions are local, thus causing the algorithm to scale very well on parallel machines.

This method is by now widely in use, for example in studies of polymers and star-polymers under shear flow[10] and vesicles or red blood cells in flow[11,12].

3 Sperm Model

Sperm cells consist of three parts. The head, containing the nucleus where the genetic information is stored, the beating tail that propels the sperm through the fluid and the midpiece in between. Of course more structures and substructures can be identified, but these are not important for their hydrodynamic properties.

The beating tail, i.e. the axoneme, can be well described as an active elastic rod. We model the axoneme with a crane-like structure (see Fig 1). Three semi-flexible rods are interconnected by springs to obtain a three-dimensional structure with well-suited elastic properties. The three-fold structure allows to impose an internal-active and directed bending by changing the rest lengths of the individual bonds, thus imposing a spontaneous directed curvature onto the structure.

The final tail structure contains 100 monomeres per rod, which are thus $S_l = 50\,a$ long. A spherical head is attached in front. For propulsion, a propagating sine-wave is imposed on the tail behind a short ($10a$) passive midpiece. In experiments, helical motion of sperm in a bulk fluid is observed, requesting some form of chirality of the sperm due to symmetry reasons. To impose this chirality, the midpiece can be bent by a predefined bending parameter b, see Fig. 4. The bending paramter b measures how much one rod of the midpiece is shortened, relative to its straight length. At $b = 9\%$ the midpiece forms approximately a quarter circle. Since this bent is imposed on a different rod than the beat pattern, it bends the head out of the beating plane, creating chirality.

Coupling to hydrodynamics is done by immersing the structure in a MPC fluid and including monomers in the collision step, as suggested in Ref.[13] for polymers. To sufficiently resolve the hydrodynamics in MPC, the beat amplitude of the tail has to exceed a

few MPC collision boxes. This requires a large sperm structure of length $S_l = 50\,a$, and correspondingly large simulation boxes. For bulk simulations, we simulate 70^3 collision boxes with periodic boundary conditions. Because each box contains 10 fluid particles, this corresponds to 3.5 million particles, and consequently demands large computational resources.

4 Hydrodynamics of Thin Rods

The tail of a sea urchin spem cell is $50\mu m$ long, and less than $0.5\mu m$ in diameter. The hydrodynamics of the tail should therefore be well described by a sequence of thin rods. Thin rods have an anisotropic friction with the solvent, where the friction force perpendicular to the rod, $\vec{F}_\perp$ is larger then parallel, $\vec{F}_\parallel$.

Intuitively it is obvious that dragging a rod perpendicular to its orientation is more difficult than parallel, and the physical origin of this behaviour is easy to understand. If moved parallel to its orientation, most of the rod can travel in the wake of the tip, thus reducing friction.

We define rod-drag coefficients for Stokes flow by

$$F_\parallel = \gamma_\parallel v_\parallel \tag{3}$$

$$F_\perp = \gamma_\perp v_\perp \tag{4}$$

with the subscript $\parallel$ for vector components parallel to the rod, and $\perp$ for perpendicular components. These friction coefficients are related to the diffusion coefficients $D_{\parallel,\perp}$ via

$$\gamma_{\parallel,\perp} = \frac{k_B T}{D_{\parallel,\perp}} \tag{5}$$

Calculating the diffusion coefficients for a rod-like colloid of finite length is not trivial, but approximations can be found in the literature. Tirado, Martinez and Garcia de la Torre[14] reviewed some theoretical approaches. The different theories agree on

$$\frac{2\pi\eta L_r D_\parallel}{k_B T} = \ln(L_r/d_r) + \nu_\parallel \tag{6}$$

$$\frac{4\pi\eta L_r D_\perp}{k_B T} = \ln(L_r/d_r) + \nu_\perp. \tag{7}$$

where L_r is the length, and d_r the diameter of the rod. Differences between theories are found concerning the correction functions $\nu_{\perp/\parallel}$. One approximation for $2 < L_r/d_r < 30$ is

$$\nu_\perp = 0.839 + 0.185 d_r/L_r + 0.233\,(d_r/L_r)^2 \tag{8}$$

$$\nu_\parallel = -0.207 + 0.980 d_r/L_r - 0.133\,(d_r/L_r)^2 \tag{9}$$

These results can be used to calculate the swimming velocity of sperm. Gray and Hancock[15] used this approach to calculate the swimming velocity of sperm cells. After assuming a sine-shaped beat-pattern, $\gamma_\perp = 2\gamma_\parallel$, and some other approximations, they obtained the swimming speed

$$\bar{v}_x = \frac{\omega\pi b^2}{\lambda}\left(1 + \frac{4\pi^2 b^2}{\lambda^2} - \sqrt{1 + \frac{2\pi^2 b^2}{\lambda^2}}\frac{C_H}{n\lambda\gamma_\parallel}\right)^{-1} \tag{10}$$

where ω is the angular frequency, λ the wavelength and b the amplitude of the beat. The number of waves present on the tail is denoted by n, C_H is the drag coefficient of the head. For a spherical head of radius R_h and a thin tail of radius d, it is possible to approximate

$$\frac{C_H}{\gamma_\parallel} = 3R_h \left[\left(\log \frac{d}{2\lambda} \right) + \frac{1}{2} \right]. \tag{11}$$

Because a pre-defined beat shape has been used, the viscosity η does not influence the swimming speed.

Like the swimming velocity of sperm, other mesoscopic systems of slender bodies can be described in this fashion. Always, either the assumption $\gamma_\perp = 2\gamma_\parallel$ or the ratio $\gamma_\perp/\gamma_\parallel$ play an essential role. Sperm, for example, would not move if $\gamma_\perp = \gamma_\parallel$. This shows the importance and effects of the anisotropic friction of thin rods. To compare our results to slender body theory, and as a suitability test of our model, we performed simulations to calculate $\gamma_\perp$ and $\gamma_\parallel$ in an MPC fluid.

For this purpose we keep the structure, presented above as the sperms tail, in the centre of a simulation box and expose it to a constant flow. The friction coefficients are determined by averaging the forces on the structure. A uniform flow field imposed sufficiently far from the rod by assigning Gaussian-distributed velocities, with an average of $\bar{v}$ in the flow direction, to the solvent particles in a layer of thicknes $1a$ perpendicular to the flow direction.

The drag force is found to be linear with $\bar{v}$ at least up to $\bar{v} \approx 0.3\sqrt{m/k_bT}$. We chose $\bar{v} = 0.1\sqrt{m/k_bT}$, well within the linear regime for the remaining simulations.

Fig. 2 shows as an example the friction coefficients $\gamma_\parallel$ and $\gamma_\perp$ as a function of scaled inverse linear system size (L_r/S). To determine the friction coefficients at infinite dilution,

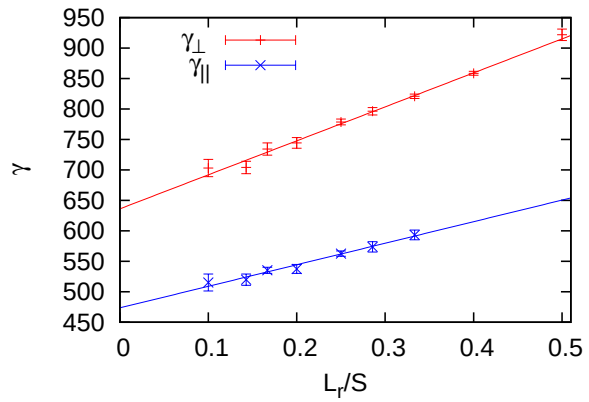


Figure 2. Friction coefficients $\gamma_\parallel$ and $\gamma_\perp$ as a function of scaled inverse linear system size L_r/S. System size S^3 varies between $(20\,a)^3$ and $(100\,a)^3$. Rod length is $10\,a$. MPC parameters are $\alpha = 130°, \rho = 10, h = 0.05\sqrt{ma^2/k_BT}$.

we fitted a linear function to the data (see Fig. 2), and used the extrapolation to infinite system size. The strong finite-size effects demand large systems and limit the range of

accessible rod lengths. The largest simulated systems used are for rods with a length of $20\,a$ with 110^3 boxes, or 13 million particles.

Considering the strong finite-size effects, and that the polymer rods are weakly penetrable for the fluid, the agreement with theory is surprisingly good (see Fig. 3). The fit in

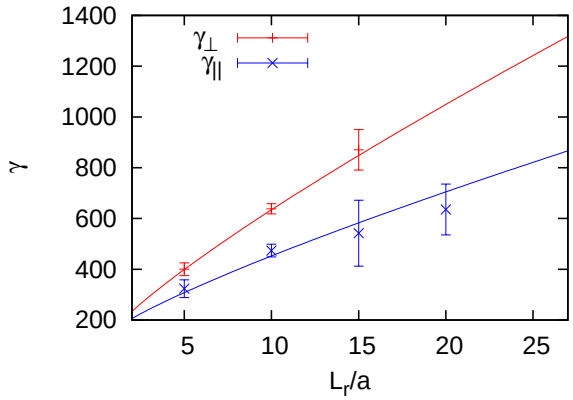


Figure 3. Friction coefficients $\gamma_{\|}$ and $\gamma_{\perp}$ as a function of rod length L_r. The solid line is a fit of the theory (Eqs.(6) and (7)) to the data of $\gamma_{\perp}$, resulting in $d_r \approx 0.9\,a$. This value was then also used to plot $\gamma_{\|}$. Points are results of finite-size fits, as described in the text.

Fig. 3 resulted in a rod diameter of $d_r \approx 0.9\,a$, which is quite reasonable. Furthermore, we see from Fig. 3 that at these low aspect ratios $\gamma_{\perp}/\gamma_{\|} < 2$. Note that also for sea urchin sperm tails, $L_r/d_r \approx 100$. The effective aspect ratio is considerably smaller due to the wave length $\lambda < L_r$.

5 Sperm Dynamics

We present here our results for bulk motion of sperm cells. Our aim is to investigate the effect of midpiece curvature and sperm chirality on the bulk motion. In bulk fluid we observe helical as well as almost straight trajectories, depending on the degree of chirality. Weakly bent sperm swim on a narrow helix. The tail is always close to the trajectory, just the head performs some sideways motions. Strongly bent sperm swim in a larger helix with the head near the centre of the helix and the tail pointing outward (see Fig. 4).

Besides its position and orientation in space, a helix has two independent parameters; typically these are the pitch p and the radius r. Additionally the sperm cells move with a velocity v on these helical trajectories. Some of these parameters are more prone to thermal fluctuations then others. We found that the curvature c, the tangential velocity v and the rotation frequency ω are the most reliable observables (see Fig. 5). Other parameters, for example the helix radius r, the velocity along the centreline v_z and the pitch p are connected to the tangential velocity v, angular frequency ω and curvature c via

$$r = \frac{cv^2}{\omega^2} \qquad v_z = \pm\sqrt{v^2 - c^2v^4} \qquad p = v_z 2\pi/\omega, \tag{12}$$

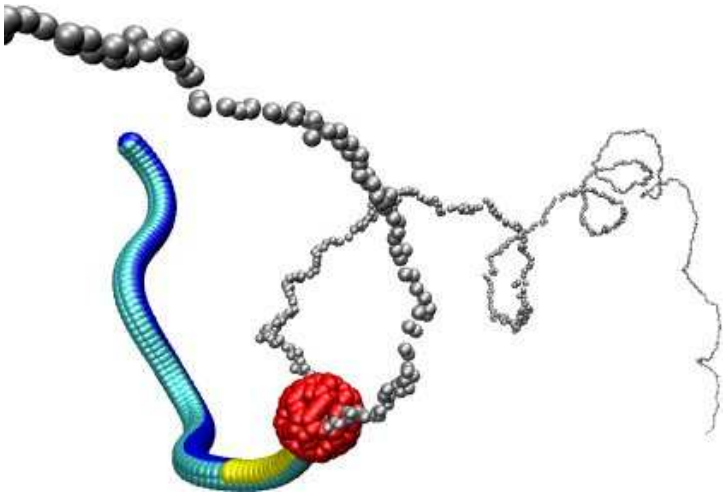


Figure 4. Visualization of a sperm motion in bulk fluid. In the beating tail the length of the blue rod is changed to impose a sinusoidal bending wave. At the tip a spherical head (red) is attached. In between we add a midpiece, that, in this case, is bend by shortening the yellow rod by 6%. At large enough bending parameters the elastic sperm twists on the axoneme, creating a strong bend. In this case, the twist is locatet approximatly at the kink in the bottom left corner of the picture. Without walls this sperm swims in a helical motion, visualized as the trajectory of the sperms head in gray spheres. Visualization using VMD[16]

and thus can be calculated from the data presented in Fig. 5.

The tangential velocity v is fairly independent of the bending parameter, with a velocity around $v \approx 0.02\, S_l/\tau_b$, decreasing by about 20% for strongly asymmetric sperm. Here τ_b is the beating period of the sperm.

The curvature follows a sigmoidal function, starting at $c < 0.5/S_l$ for the symmetric sperm, increasing rapidly at $b \approx 4\%$ to curvatures around $2/S_l$. The trajectory looks helical if the average curvature is larger then $1/S_l$, while curvatures smaller then $1/S_l$ are generally found if the trajectory is rather straight. Stronger curvature is found for sperm with a significant bent in the tail. This bent is generated dynamically for $b > 4\%$. In this case, the thrust of the tail and the drag force of the head induce the twisting and bending of the axoneme out of the beating plane. With careful observation this twist can also be seen in Fig. 4.

This rapid change in the tail curvature also effects other swimming parameters, like the rotation frequency. The rotation frequency at first increases with the bending parameter, but as the sperm is deformed, rotation slows down. The elastic deformation of the tail increases its asymmetry, so that the head bends even further out of the beating plane, thereby slowing down the rotational motion.

6 Outlook

The natural extension of this work is to look at hydrodynamic interactions of active objects. The first such question is what happens in the presence of walls. It is known for quite some time[17] that sperm cells accumulate at the walls of an observation chamber. Preliminary

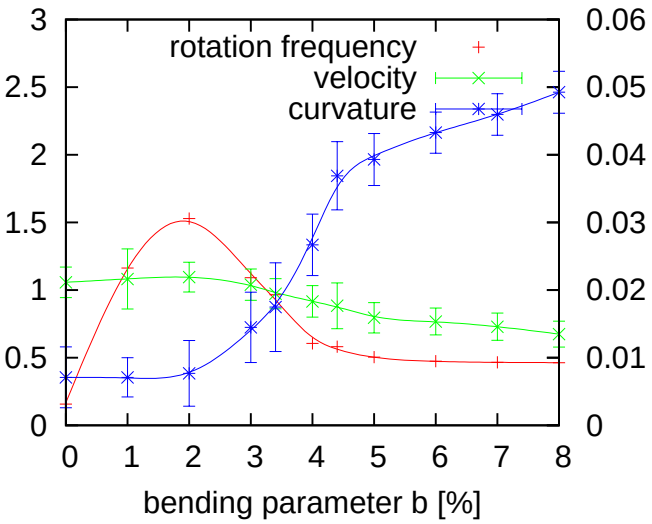


Figure 5. Parameters of helical sperm trajectories as a function of bending parameter b. All quantities are measured in units of sperm length S_l and beat period τ_b. Curvature is given in units of $1/S_l$ (left scale), rotation frequency in units of $2\pi/\tau_b$ (right scale), and swimming velocity in units of S_l/τ_b (right scale)

simulation results show that hydrodynamic interactions are sufficent to cause an effective attraction of sperm cells to walls.

The true strength of this simulation scheme is the possibility to simulate several interacting active objects. Large arrays of biological cilia display metachronal waves. We have simulated large arrays of cilia and have shown that hydrodynamics is sufficient to cause a self-organized metachronal coupling. Furthermore we could show that this metachronal wave enhances fluid transport and efficiency by almost an order of magnitude.

Acknowledgments

We thank NIC for providing computational resources on JUMP. Helpful discussions with U.B. Kaupp are gratefully acknowledged.

References

1. E. M. Purcell. Life at low Reynolds-number. *Am. J. Phys.*, **45**:3–11, 1977.
2. `http://physics.nyu.edu/pine/research/hydroreverse.html`.
3. T. Strunker, I. Weyand, W. Bonigk, Q. Van, A. Loogen, J. E. Brown, N. Kashikar, V. Hagen, E. Krause, and U. B. Kaupp. A K+-selective cGMP-gated ion channel controls chemosensation of sperm. *Nature Cell Biol.*, **8**:1149, 2006.
4. A. Malevanets and R. Kapral. Mesoscopic model for solvent dynamics. *J. Chem. Phys.*, **110**:8605–8613, 1999.

5. A. Lamura, G. Gompper, T. Ihle, and D. M. Kroll. Multi-particle collision dynamics: Flow around a circular and a square cylinder. *Europhys. Lett.*, **56**:319–325, 2001.
6. T. Ihle and D. M. Kroll. Stochastic rotation dynamics: A Galilean-invariant mesoscopic model for fluid flow. *Phys. Rev. E*, **63**:020201, 2001.
7. M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper. Low-Reynolds-number hydrodynamics of complex fluids by multi-particle collision dynamics. *Europhys. Lett.*, **68**:106–112, 2004.
8. M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper. Dynamic regimes of fluids simulated by multiparticle-collision dynamics. *Phys. Rev. E*, **72**:016701, 2005.
9. T. Ihle and D. M. Kroll. Stochastic rotation dynamics. I. Formalism, Galilean invariance, and Green-Kubo relations.
10. M. Ripoll, R. G. Winkler, and G. Gompper. Star polymers in shear flow. *Phys. Rev. Lett.*, **96**:188302, 2006.
11. H. Noguchi and G. Gompper. Fluid vesicles with viscous membranes in shear flow. *Phys. Rev. Lett.*, **93**:258102, 2004.
12. H. Noguchi and G. Gompper. Shape transitions of fluid vesicles and red blood cells in capillary flows. *P. Natl. Acad. Sci. USA.*, **102**:14159, 2005.
13. A. Malevanets and J. M. Yeomans. Dynamics of short polymer chains in solution. *Europhys. Lett.*, **52**:231–237, 2000.
14. M. M. Tirado, C. L. Martinez, and J. G. Delatorre. Comparison of theories for the translational and rotational diffusion-coefficients of rod-like macromolecules - application to short DNA fragments. *J. Chem. Phys.*, **81**:2047–2052, 1984.
15. J. Gray and G. J. Hancock. The propulsion of sea-urchin spermatozoa. *J. Exp. Biol.*, **32**:802–814, 1955.
    Phys. Rev. E, **67**:066705, 2003.
16. W. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, **14**:33–38, 1996.
17. L. Rothschild. Non-random distribution of bull spermatozoa in a drop of sperm suspension. *Nature*, **198**:1221, 1963.

Optimization of a Physics-Based United-Residue Force Field (UNRES) for Protein Folding Simulations

Adam Liwo[1,3], Cezary Czaplewski[1,3], Stanisław Ołdziej[2,3], Urszula Kozłowska[3], Mariusz Makowski[1,3], Sebastian Kalinowski[1], Rajmund Kaźmierkiewicz[2,3], Hujun Shen[3], Gia Maisuradze[3], and Harold A. Scheraga[3]

[1] Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland
E-mail: {*adam, czarek, momo, bes*}*@chem.univ.gda.pl*

[2] Intercollegiate Faculty of Biotechnology, University of Gdańsk, Medical University of Gdańsk
ul. Kładki 24, 80-822 Gdańsk, Poland
E-mail: {*stan, rajmund*}*@biotech.ug.gda.pl*

[3] Baker Laboratory of Chemistry and Chemical Biology, Cornell University
Ithaca, NY 14853-1301, U.S.A.
E-mail: {*uad2, hs322, gm56, has5*}*@cornell.edu*

Understanding the functioning of living cells requires knowledge of structure and long-time dynamics of proteins and other biological macromolecules, which information is not readily available from experiment. The development of distributed computing has opened new avenues for such studies. Further, reduction of the representation of polypeptide chains to the so-called united-residue or coarse-grained representation enables the extension of the time scale of calculations to micro- or even milliseconds. In this report, we describe recent developments of the united-residue (UNRES) force field for large-scale simulations of protein structures and dynamics carried out with the use of the resources at the Supercomputer Centre in Jülich.

1 Introduction

One of the major and still unsolved problems of computational biology is to understand how interatomic forces determine how proteins fold into the three-dimensional structures. The practical aspect of the research on this problem is to design a reliable algorithm for the prediction of the three-dimensional structure of a protein from its amino-acid sequence, which is of utmost importance because experimental methods for determination of protein structures cover only about 10% of new protein sequences. In the case of protein structure prediction, methods that implement direct information from structural data bases (e.g., homology modelling and threading) are, to date, more successful compared to physics-based methods[1]; however only the latter will enable us to extend the application to simulate protein folding and to understand the folding and structure-formation process. The underlying principle of physics-based methods is the *thermodynamic hypothesis* formulated by Anfinsen[2], according to which the ensemble called the "native structure" of a protein constitutes the basin with the lowest free energy under given conditions. Thus, energy-based protein structure prediction is formulated in terms of a search for the basin with the lowest free energy; in a simpler approach the task is defined as searching for the conformation with the lowest potential energy[3], and prediction of the folding pathways can be formulated as a search for the family of minimum-action pathways leading to this basin from the unfolded (denaturated) state. In neither procedure do we want to make use of ancillary data

from protein structural databases. Equally important is to simulate the pathways of protein folding, misfolding (which is the cause of prion diseases, cancer, and amyloid diseases) and large-scale conformational changes which occur during enzymatic catalysis or signal transduction.

United-residue (also termed coarse-grained or mesoscopic) representations of polypeptide chains enable us to carry out large-scale simulations of protein folding, to study protein free-energy landscapes and to carry out physics-based predictions of protein structure[4]. Owing to the considerable reduction of the number of interacting sites and variables, the cost of computations is reduced hundreds or thousand of times compared to an all-atom representation of polypeptide chains in implicit and explicit solvent, respectively; this enables micro- or even millisecond simulations of protein folding to be carried out. For the past several years, we have been developing a physics-based coarse-grained UNRES model of polypeptide chains and the corresponding force field[5–14]. Initially[7], it was designed for physics-based predictions of protein structure through global optimization of an effective potential-energy function of polypeptide chains plus solvent. With this approach, we achieved considerable success in blind-prediction CASP exercises[12]. Recently[15] we implemented a mesoscopic dynamics method to the UNRES force field which enabled us to carry out real-time *ab initio* simulations of protein folding. Subsequently, we implemented[16] the replica-exchange (REMD)[17] and multiplexing-replica exchange (MREMD)[18] extensions of MD, which enabled us to study the thermodynamics of protein folding. However, the new applications required reparameterization of the UNRES force field to reproduce the thermodynamic characteristics of protein folding.

2 Methods

2.1 The UNRES Model of Polypeptide Chains

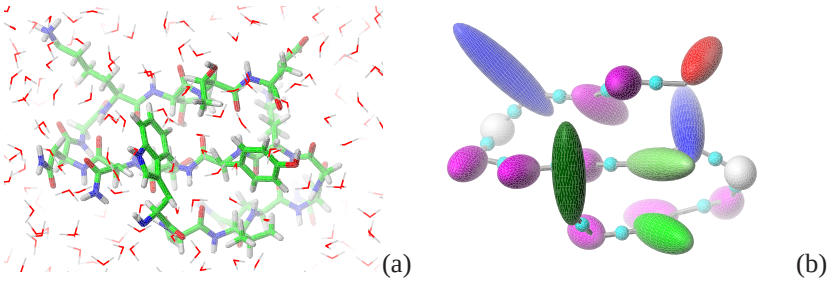


Figure 1. Illustration of the correspondence between the all-atom polypeptide chain in water (a) and its UNRES representation (b). The side chains in part (b) are represented by ellipsoids of revolution and the peptide groups are represented by small spheres in the middle between consecutive α-carbon atoms. The solvent is implicit in the UNRES model.

In the UNRES model[5–14], a polypeptide chain is represented by a sequence of α-carbon (C^{α}) atoms linked by virtual bonds with attached united side chains (SC) and united peptide groups (p). Each united peptide group is located in the middle between two consecutive

α-carbons. Only these united peptide groups and the united side chains serve as interaction sites, the α-carbons serving only to define the chain geometry, as shown in Figure 1. The $C^{\alpha} \cdots C^{\alpha}$ virtual bond lenghts (i.e., the distances between neighbouring C^{α}'s) are 3.8 Å corresponding to *trans* peptide groups.

The effective energy function is a sum of different terms corresponding to interactions between the SC ($U_{SC_iSC_j}$), SC and p ($U_{SC_ip_j}$), and p ($U_{p_ip_j}$) sites, as well as local terms corresponding to bending of virtual-bond angles θ (U_b), side-chain rotamers (U_{rot}), virtual-bond torsional (U_{tor}) and double-torsional (U_{tord}) terms, virtual-bond-stretching (U_{bond}) terms, correlation terms ($U_{corr}^{(m)}$) pertaining to coupling between backbone-local and backbone-electrostatic interactions[8] (where m denotes the order of correlation), and a term accounting for the energetics of disulfide bonds (U_{SS}). Each of these terms is multiplied by an appropriate weight, w, which must be determined by optimization of the energy function by using training proteins. The energy function is given by equation 1.

$$\begin{aligned} U = & w_{SC} \sum_{i<j} U_{SC_iSC_j} + w_{SCp} \sum_{i \neq j} U_{SC_ip_j} + w_{pp} \sum_{i<j-1} U_{p_ip_j} \\ & + w_{tor} \sum_i U_{tor}(\gamma_i) + w_{tord} \sum_i U_{tord}(\gamma_i, \gamma_{i+1}) \\ & + w_b \sum_i U_b(\theta_i) + w_{rot} \sum_i U_{rot}(\alpha_{SC_i}, \beta_{SC_i}) + \sum_{m=3}^{6} w_{corr}^{(m)} U_{corr}^{(m)} \\ & + w_{bond} \sum_{i=1}^{nbond} U_{bond}(d_i) + w_{SS} \sum_i U_{SS;\, i} \end{aligned} \tag{1}$$

The method of optimizing the force field developed in our laboratory is termed *hierarchical method*[11,14] and aims at obtaining such energy landscapes of selected training proteins that the free energy of each of the training proteins decreases with increasing native likeness. The conformational space is discretized into levels, each of which corresponds to a certain degree of native likeness. In the present study we computed the free energies below, at, and above the folding-transition temperatures and extended the approach by the requirements that the free-energy relations be inverted above the folding-transition temperature. This is illustrated in Figure 2.

Given a set of training proteins, optimization of the force field consists of iterating the cycles, each consisting of (i) simulations with current parameters of the energy function, (ii) computing the free energies at selected temperatures from the simulation data, and (iii) adjusting the parameters of the energy function to achieve the desired relations between the free energies of the sub-ensembles of the training proteins (Figure 2). The procedure is terminated when the required relations between free energies hold after a new simulation with optimized parameters. The primary optimizable parameters were the energy-term weights of equation 1. To generate the decoy sets for optimization at various temperatures simultaneously, we implemented the multiplexing replica-exchange molecular dynamics (MREMD)[17,18] in UNRES[16]. We developed a parallel well-scalable code for the UNRES/MREMD method, which scales 75% up to 4096 processors (Figure 3). To compute free energies and other ensemble-related quantities from simulation data, we implemented the weighted-histogram analysis method (WHAM)[19].

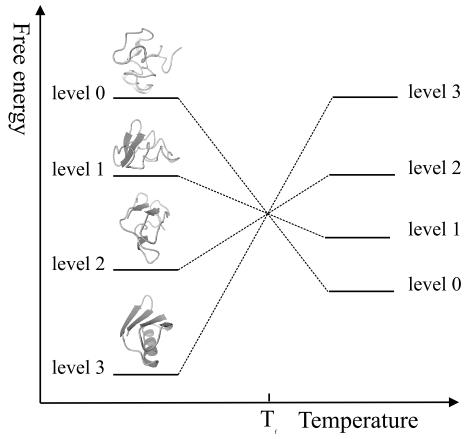


Figure 2. Illustration of ordering of the energy levels, which is the goal of the algorithm for optimizing the potential function[14], using the 1EM7 protein of the IGG family, the structure of which consists of two β-hairpins packed to a middle α-helix. Only one conformation has been selected to represent each of the structural levels. Below the folding-transition temperature (T_f), the non-native level (level 0) has the highest free energy, the conformations with only the native C-terminal β-hairpin forming (level 1) have a lower free energy, next are the conformations in which the middle part of the N-terminal β-strand joins the β-hairpin and the middle α-helix starts to form and, finally, the native-like structures with all structural elements formed have the lowest free energy. Above the folding-transition temperature the free-energy relations are reversed and at the folding-transition temperature the free energies should be approximately equal.

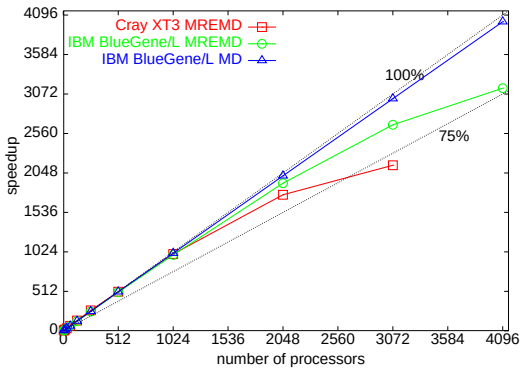


Figure 3. Speedup plots for the MREMD code using IBM Blue Gene/L (green circles) and Cray XT3 (red rectangles). For comparison, data from an ideally scalable series of independent canonical MD runs are shown. The system is the 1SAP protein.

3 Results

Initially[14] we applied the new optimization procedure to three proteins separately identified by the following PDB codes: 1E0L (a 28-residue anti-parallel three-stranded β-sheet), 1GAB (a 47-residue three-α-helix bundle), and 1E0G (a 48-residue $\alpha + \beta$ protein). All three force fields exhibited a heat-capacity peak corresponding to the folding-transition

temperature. The force field optimized on 1GAB was fairly transferable to other α-helical proteins, whose native-like ensembles of structures were located within the five most probable clusters of structures[14].

As the next step, we carried out hierarchical optimization using two training proteins: 1ENH (a three-helix bundle; 56 residues) and the full 37-residue sequence of 1E0L. This choice was motivated by the availability of the experimental temperature dependence of the free energy[20,21] and by the fact that these proteins, although small in size, contain significant regions with undefined secondary structure, which are hard to reproduce in simulations. To generate a converged statistical ensemble of conformations for one protein, typically 20,000,000 MREMD steps with 1024 processors are required, which means 1 rack-week of computation with Blue Gene/L. The total computational effort to optimize the force field on 1E0L and 1ENH was 2 rack-months with Blue Gene/L.

The most probable conformations of the two training proteins calculated with the optimized parameters of the UNRES energy function at room temperature are superposed on the experimental structure on Figure 4. The experimental[20,21] and calculated free-energy gaps vs. temperature are compared in Figure 5, while the calculated heat-capacity and RMSD curves vs. temperature are shown in Figure 6.

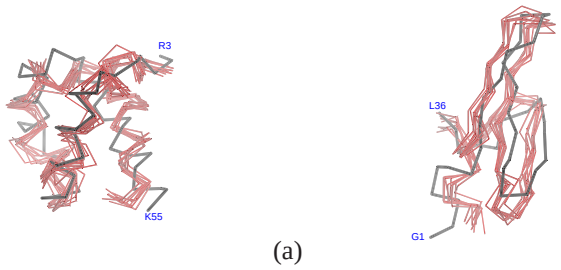


Figure 4. The C^{α} traces of 10 most probable conformations at T=300°K of 1ENH (a) and 1E0L (b) calculated with the UNRES force field optimized on these two proteins (red lines) superposed on the C^{α} traces of the corresponding experimental structures (black lines)[20,21].

With the optimized force field, we carried out MREMD simulations of the mutants of 1E0L studied by Gruebele et al.[20] (these mutations result in a shift of the folding temperature). Subsequently, we calculated the heat-capacity curves of 1E0L mutants and determined their folding temperatures. All mutants studied except for the W30A mutant folded to structures similar to that of the wild-type protein (with ensemble-averaged RMSD at room temperature from 3.5 to 4.5 Å). The calculated folding temperatures are compared with their experimental counterparts in Table 1.

4 Conclusions

The results of our research demonstrated that it is possible to obtain a coarse-grained force field for protein simulations which reproduces thermodynamic properties of wild-type proteins and their mutants, as well as is transferable to proteins outside the training set. Consequently, large-scale simulations of protein structure and dynamics are at hand. Our current

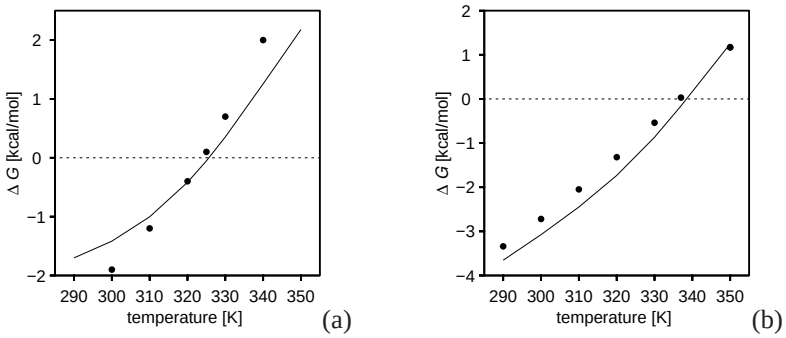


Figure 5. Calculated with optimized force field (lines) and experimentally determined (filled circles) free energy of folding of (a) 1ENH and (b) 1E0L.

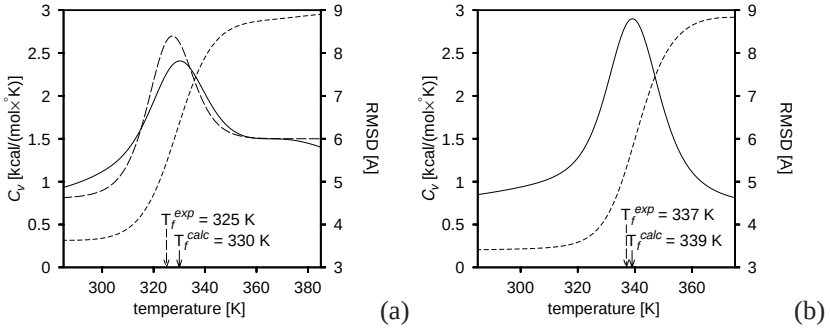


Figure 6. Calculated heat-capacity curves (solid lines) and RMSD curves vs. temperature (short-dashed lines) of (a) 1ENH and (b) 1E0L. The long-dashed curve shown in panel (a) is the experimental heat-capacity curve of 1ENH shifted vertically to match the tail of the calculated heat-capacity curves. The calculated and experimental folding temperatures are also shown. The experimental heat-capacity curve of 1E0L has not been determined.

research is focused on improving the parameterization of UNRES and using the force field to study the kinetics of protein folding, protein aggregation, and large-scale motions.

Acknowledgments

This research was conducted by using the resources of the John von Neumann Institute for Computing at the Central Institute for Applied Mathematics, Forschungszentrum Jülich, Germany and was financially supported by grants from the National Institutes of Health (GM-14312), the National Science Foundation (MCB05-41633), the NIH Fogarty International Center (TW7193), and grant DS 8372-4-0138-7 from the Polish Ministry of Science and Higher Education.

Mutant	T_f^{calc}	T_f^{exp}
WT	339	337
W30F	334	339
W30A	–[a]	328
Y11R	342	339
Y19L	319	328
DNY11R[b]	335	339
DNDCY11R[c]	325	328

Table 1. Experimental and calculated, with the optimized force field, folding temperatures of mutants of 1E0L.

[a]This mutant did not fold in MREMD simulations.
[b]Deletion of the 6-residue N-terminal fragment.
[c]Deletion of the N-terminal and the C-terminal fragments.

References

1. J. Moult. Rigorous performance evaluation in protein structure modelling and implications for computational biology. *Phil. Trans. R. Soc. B*, 361:453–458, 2006.
2. C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181:223–230, 1973.
3. H. A. Scheraga, J. Lee, J. Pillardy, Y.-J. Ye, A. Liwo, and D. R. Ripoll. Surmounting the multiple-minima problem in protein folding. *J. Global Optimization*, 15:235–260, 1999.
4. S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein. Coarse grain models and the computer simulations of soft materials. *J. Phys. Condens. Matter*, 16:R481–R512, 2004.
5. A. Liwo, S. Ołdziej, M. R. Pincus, R. J. Wawak, S. Rackovsky, and H. A. Scheraga. A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J. Comput. Chem.*, 18:849–873, 1997.
6. A. Liwo, M. R. Pincus, R. J. Wawak, S. Rackovsky, S. Ołdziej, and H. A. Scheraga. A united-residue force field for off-lattice protein-structure simulations. II: Parameterization of local interactions and determination of the weights of energy terms by Z-score optimization. *J. Comput. Chem.*, 18:874–887, 1997.
7. A. Liwo, J. Lee, D. R. Ripoll, J. Pillardy, and H. A. Scheraga. Protein structure prediction by global optimization of a potential energy function. *Proc. Natl. Acad. Sci., U. S. A.*, 96:5482–5485, 1999.
8. A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga. Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. *J. Chem. Phys.*, 115:2323–2347, 2001.
9. S. Ołdziej, U. Kozłowska, A. Liwo, and H. A. Scheraga. Determination of the potentials of mean force for rotation about $C^\alpha \cdots C^\alpha$ virtual bonds in polypeptides from the *ab initio* energy surfaces of terminally-blocked glycine, alanine, and proline. *J. Phys. Chem. A*, 107:8035–8046, 2003.

10. A. Liwo, S. Ołdziej, C. Czaplewski, U. Kozłowska, and H. A. Scheraga. Parameterization of backbone-electrostatic and multibody contributions to the UNRES force field for protein-structure prediction from *ab initio* energy surfaces of model systems. *J. Phys. Chem. B*, 108:9421–9438, 2004.
11. S. Ołdziej, A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga. Optimization of the UNRES force field by hierarchical design of the potential-energy landscape: II. Off-lattice tests of the method with single proteins. *J. Phys. Chem. B*, 108:16934–16949, 2004.
12. S. Ołdziej, C. Czaplewski, A. Liwo, M. Chinchio, M. Nanias, J. A. Vila, M. Khalili, Y. A. Arnautova, A. Jagielska, M. Makowski, H. D. Schafroth, R. Kaźmierkiewicz, D. R. Ripoll, J. Pillardy, J. A. Saunders, Y.-K. Kang, K. D. Gibson, and H. A. Scheraga. Physics-based protein-structure prediction using a hierarchical protocol based on the UNRES force field – test with CASP5 and CASP6 targets. *Proc. Natl. Acad. Sci. U.S.A.*, 102:7547–7552, 2005.
13. M. Khalili, A. Liwo, F. Rakowski, P. Grochowski, and H. A. Scheraga. Molecular dynamics with the united-residue (UNRES) model of polypeptide chains. I. Lagrange equations of motion and tests of numerical stability in the microcanonical mode. *J. Phys. Chem. B*, 109:13785–13797, 2005.
14. A. Liwo, M. Khalili, C. Czaplewski, S. Kalinowski, S. Ołdziej, K. Wachucik, and H.A. Scheraga. Modification and optimization of the united-residue (UNRES) potential energy function for canonical simulations. I. Temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *J. Phys. Chem. B*, 111:260–285, 2007.
15. A. Liwo, M. Khalili, and H. A. Scheraga. Molecular dynamics with the united-residue (UNRES) model of polypeptide chains; test of the approach on model proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 102:2362–2367, 2005.
16. M. Nanias, C. Czaplewski, and H. A. Scheraga. Replica exchange and multicanonical algorithms with the coarse-grained united-residue (UNRES) force field. *J. Chem. Theor. Comput.*, 2:513–528, 2006.
17. U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger. Molecular dynamics, Langevin and hybrid Monte Carlo simulations in multicanonical ensemble. *Chem. Phys. Lett.*, 259:321–330, 1996.
18. Y. M. Rhee and V. S. Pande. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophys J.*, 84:775–786, 2003.
19. S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. The weighted histogram analysis method for free-energy calculations of biomolecules. I. The method. *J. Comput. Chem.*, 13:1011–1021, 1992.
20. H. Nguyen, M. Jäger, A. Moretto, M. Gruebele, and J. W. Kelly. Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc. Natl. Acad. Sci. U.S.A.*, 100:3948–3953, 2003.
21. U. Mayor, J. G. Grossman, N. W. Foster, S. M. V. Freund, and A. R. Fersht. The denaturated state of engrailed homeodomain under denaturating and native conditions. *J. Mol. Biol.*, 333:977–991, 2003.

Transition Metal Centers in Biological Matrices: Why Nature Has Chosen Vanadate as Cofactor for Haloperoxidase

Daniel Geibig, Rainer Wilcken, Masroor Bangesh, and Winfried Plass

Institut für Anorganische und Analytische Chemie, Lehrstuhl für Anorganische Chemie II
Friedrich-Schiller-Universität Jena, 07743 Jena, Germany
E-mail: sekr.plass@uni-jena.de

While molybdate(VI) shows superior catalytic activity towards halide oxidation in solution as compared to its close chemical relative vanadate(V), the vanadium haloperoxidase enzymes nevertheless employ the latter as cofactor. How can the enormous effectiveness of the enzyme be explained? What are the details of the reaction mechanism? What are the chemical features differentiating relevant vanadium and molybdenum species? These are the questions that we posed in our computational DFT studies on several model systems of haloperoxidases. We are able to postulate an enzymatic mechanism and to explain why there are no molybdenum haloperoxidase enzymes.

1 Introduction and Motivation

The need to save energy wherever possible dominates today's life. In chemistry, the efforts in this direction have been summarized by the term "Green Chemistry". In order to produce basic chemical compounds, processes consuming a lot of energy are often necessary. Using catalysts, energy consumption can be lowered by a great amount. However, developing new catalytic compounds is by no means an easy task.

Nature uses very effective and complex catalysts – the enzymes. Imitating nature in order to develop catalysts working in a similar way to enzymes is thus an important research area of both chemists and biochemists alike. To achieve this goal, it is necessary to first elucidate the underlying catalytic mechanism. Using computational chemistry and employing suitable model systems, it is possible to study chemical reactions at the molecular level. In this project, the enzyme family of vanadium haloperoxidases (VHPOs) is investigated.

This group of enzymes is found in several seaweeds as well as fungi and plays an important role synthesizing halogenated organic compounds in natural environments. VHPOs catalyze the two-electron oxidation of a halide ion by hydrogen peroxide yielding hypohalous acid given in Eq. 1, which in turn is able to halogenate various organic compounds.

$$\mathrm{H_2O_2} + \mathrm{H^+} + \mathrm{X^-} \rightarrow \mathrm{H_2O} + \mathrm{HOX} \quad (1)$$

The structure of the active site in vanadium chloroperoxidase from *curvularia inaequalis* was investigated by X-ray crystallography and published by Wever et al. in 1996[1]. In subsequent years several studies were conducted to further elucidate the mechanism of the catalytic cycle[2–6]. The centre of the active site consists of a vanadate(V) ion in trigonal bipyramidal geometry, bound to His_{496} and forming hydrogen bonds to the surrounding amino acids[7] (Fig. 1). By a reaction with hydrogen peroxide and elimination of the re-

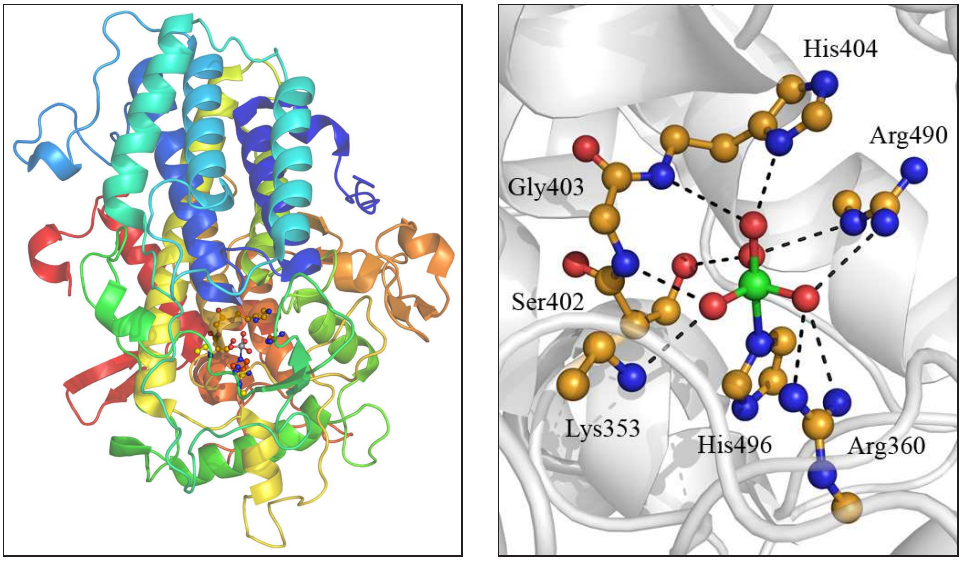


Figure 1. Left: Crystal structure of native vanadium-containing chloroperoxidase from *curvularia inaequalis* (PDB ID: 1IDQ)[3]. Right: Cutout of the active site.

sulting water molecule, a reactive peroxo species is formed (Fig. 2). Several studies have indicated that the peroxo group is activated by protonation before oxidizing the halide ion in a second step[8]. Then, the halide ion attacks the reactive oxygen atom in a nucleophilic manner and a hypohalous acid or similar "X^+" species is formed.

Due to the fact that these compounds possess versatile applicability in reactions with different nucleophilic acceptors, it is desirable to develop catalysts mimicking haloperoxidases. The simplest concept would be to employ solely the cofactor vanadate(V)[9]. However, it was found that molybdate(VI) is actually 45 times more reactive than vanadate(V), a surprising fact that was explained by the ability of molybdate(VI) to form reactive oxodiperoxo species[10–12]. Vanadate(V) is also able to form oxodiperoxo species which nevertheless are not able to oxidize bromide ions[13].

A setback to the concept of using molybdate(VI) instead of vanadate(V) as cofactor came with H. Vilter's reconstitution experiments using the apoprotein of VHPO from *ascophyllum nodosum*[14]. Replacing the vanadate cofactor with molybdate results in a total loss of catalytic activity. Taking the very similar ionic radii of V^{5+} and Mo^{6+} into account, it can however be suggested that the molybdenum cation can take the position of the vanadium cation in the enzyme.

The resulting question is: Why does the holoenzyme of vanadium bromoperoxidase not show catalytic activity when employing a molybdate cofactor, when in contrast molybdate(VI) possesses catalytic activity in solution that by far surpasses the activity of vanadate(V)? All previously made observations indicate that the reason might be differences in the catalytic mechanisms of both species. Investigations of these differences are one important goal of our work.

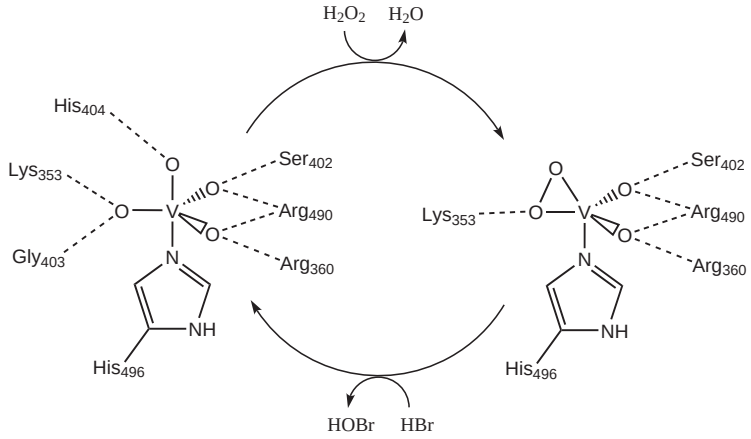


Figure 2. Scheme of the enzymatic cycle, showing the native form of the vanadium-containing chloroperoxidase (left) as well as the reactive peroxo species (right) according Ref. 3.

2 Computational Methods

It is common knowledge in quantum chemistry that density functional theory according to the Kohn-Sham scheme is well suited to investigate transition metal complexes. All DFT calculations in this work were performed using the program packages *Amsterdam Density Functional*[15–17] (ADF) and TURBOMOLE[18]. Due to the fact that the Schrödinger equation cannot be solved exactly for many-electron molecular systems, their quantum chemical treatment requires several approximations. The quality of a DFT calculation mainly depends on two factors, the nature of the exchange-correlation functional E_{xc} employed and the basis set. In this work, the gradient corrected (GGA) functional of Becke[19] and Perdew[20] (BP86) was used together with the TZVP basis set which is of split-valence triple-ζ quality with additional polarization functions on all atoms. In calculations with the ADF package, the scalar ZORA (*zero-order regular approximation*) method implemented into ADF by E. van Lenthe et al.[21] was used to account for relativistic effects of molybdenum. In the TURBOMOLE calculations, a relativistic ECP (*effective core potential*)[22] for molybdenum was used.

Massively parallelized high performance computer systems are the most important tool for computer chemists, as larger basis sets and more complex functionals can be employed in order to reach more accurate results.

Carrying out geometry optimizations of molecular structures, stationary points such as local minima and maxima on the many-dimensional potential surface can be located. A verification of the nature of intermediate structures (local minima), such as reactants and products, and transition state structures (local maxima) can be achieved by means of vibrational frequency calculations. These frequency calculations also yield thermodynamic data such as enthalpy and entropy values which allow for the construction of a Gibbs free energy profile for all reaction steps, including several competing reaction paths. The path possessing the lowest energy is the most probable path for the reaction. Nevertheless, it is

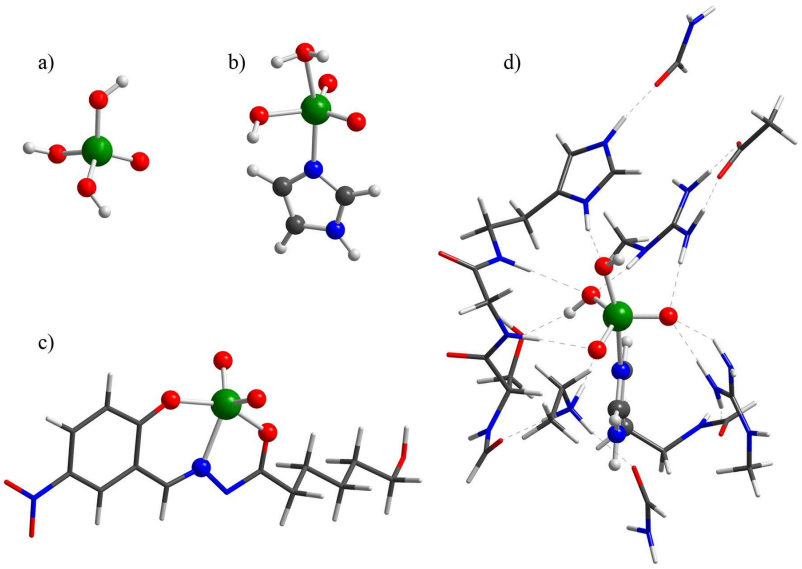


Figure 3. Several structures of model systems for investigations concerning haloperoxidase activity, optimized with TURBOMOLE/BP86(RI)/TZVP. a) vanadate(V), b) imidazole-bound vanadate(V) with apical water, c) example vanadium complex $[VO_2(NO_2salhyhh)]^-$, d) cutout of the active site of VCPO.

necessary to investigate every chemically possible reaction path, a process which consumes a lot of computational time. Using quantum chemistry, it is thus possible to understand and explain mechanisms of chemical reactions.

3 Molecular Model Systems

Even with today's arrays of high performance computer clusters, it is still impossible to treat large molecular systems like whole enzymes in a quantum chemical way. Thus molecular model systems represent a central tool to elucidate chemical reaction mechanisms. The choice of a model system has decisive influence on the results obtained. In general it can be stated that larger model systems lead to more accurate results than smaller ones. In Fig. 3 different model systems whose haloperoxidase activity was investigated in this project are shown. In addition to the vanadium species shown, the corresponding molybdenum species were also examined to uncover differences in the catalytic cycles. Additional systems with varying degrees of protonation can also be envisaged. Structures a) to c) are very simple model systems sharing only basic similarity with the active site of VHPO. Their advantage lies in the fact that they are made up of a small number of atoms, requiring less computational time than system d) and thus allowing the investigation of numerous possible reaction paths. Interpreting the results obtained in a reasonable way, plenty of information regarding the catalytic cycle of the enzyme can be obtained.

Model system d) represents a cutout of the active site of vanadium chloroperoxidase. Solely in this model system, the important influence of the hydrogen bonds formed by vanadate and the surrounding amino acid is considered. It thus constitutes the best model

system but introduces additional problems, apart from the dramatically increased computational demands: The positions of a total of 26 atoms of the protein backbone must be fixed in order to prevent the molecular system – which is now not held together anymore by the enzyme matrix – from falling apart and to preserve the active site structure.

4 Catalytic Mechanism of the Enzyme

Many open questions in connection with the catalytic mechanism of the VHPOs were answered by extensive theoretical studies in recent years, and finally an enzymatic cycle, shown in Fig. 4, has been postulated[5,23]. As can easily be seen, the amino acids bound to the cofactor by hydrogen bonds play an important role. In the first reaction step, the protonated His_{404} delivers its proton to the axial hydroxo group of the cofactor, forming a water molecule (4b). This water molecule can now be replaced by hydrogen peroxide in a dissociative manner, which then leads to the catalytically active peroxo species after the elimination of an additional water molecule (4c-d). In the second step of the catalytic reaction, Lys_{353} activates the peroxo group of the cofactor by means of hydrogen bonding, in this way facilitating the halide oxidation process (4e-f).

These results were obtained using the model system b) of Fig. 3 selectively expanded by individual amino acid residues. Due to limited computational resources, only selected

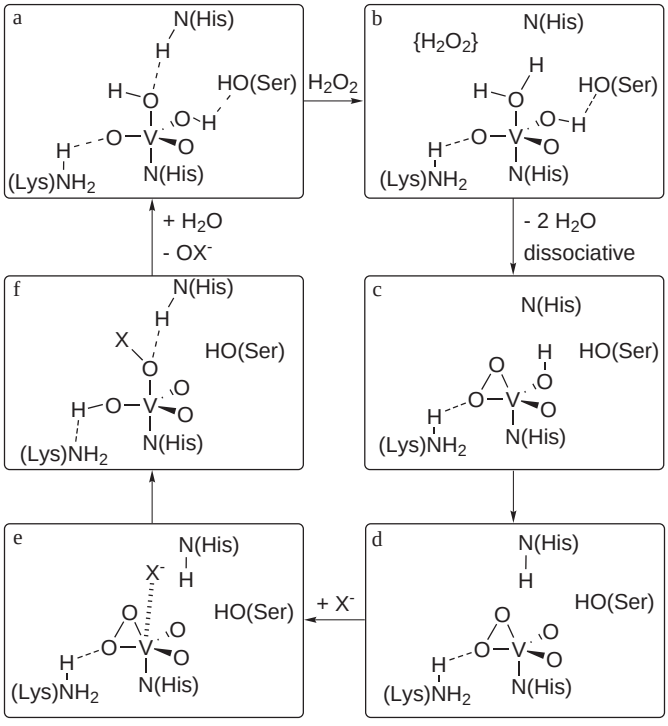


Figure 4. Proposed catalytic cycle for the halide oxidation by vanadium-containing haloperoxidases[23]; emphasizing the specific role of the amino acid residues Lys_{353}, Ser_{402} and His_{404} (see Fig. 1).

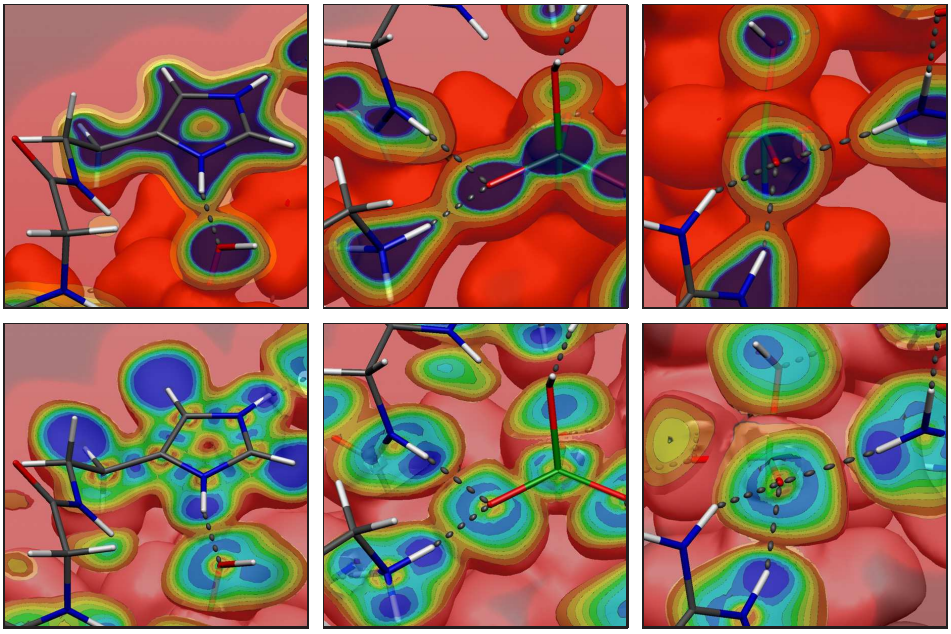


Figure 5. Plot of the electron density (first row) as well as a plot of the electron localization function (second row), highlighting selected hydrogen bonds. From left to right: His_{404}–$\mathrm{H}\cdots\mathrm{OH}_{ax}$, Lys_{353}–$\mathrm{H}\cdots\mathrm{O}_{eq}$ and Arg_{360}–$\mathrm{H}\cdots\mathrm{O}_{eq}$.

results could be verified using model system d). Investigations of the hydrogen bonding network in the native structure of the active site (model system d) by electron density and electron localization function (ELF) calculations (shown in Fig. 5) yielded particularly interesting results. Emphasis in further computational studies will be laid on an investigation of changes in the hydrogen bonding network during the course of the catalytic reaction.

5 Difference between Vanadium and Molybdenum Species

In order to understand why nature prefers vanadate(V) instead of choosing the catalytically more active molybdate(VI) as cofactor for haloperoxidases, DFT calculations on molybdate model systems similar to b) in Fig. 3 were performed. The results obtained show that the oxodiperoxo species $[\mathrm{MoO(O_2)_2Im}]$ is indeed more catalytically active than the dioxo-monoperoxo species $[\mathrm{MoO_2(O_2)Im}]$. However, the spatial environment of the VHPO active site is tailored to a monoperoxo species. Calculations investigating the halide oxidation activity of the molybdenum(VI) monoperoxo species allow the conclusion that $[\mathrm{MoO_2(O_2)Im}]$ exhibits some catalytic activity – a contradiction to the results obtained by Vilter's studies[14] of the VHPO holoprotein. Presumably, the actual reason for the inactivity of "molybdenum bromoperoxidase" might be that no peroxo species is formed in the first place.

Ongoing studies hence explore the formation of the molybdenum monoperoxo species. The first and central step of the dissociative mechanism is the cleavage of water (Fig. 4b) and the subsequent attack by hydrogen peroxide. Since molybdenum generally supports

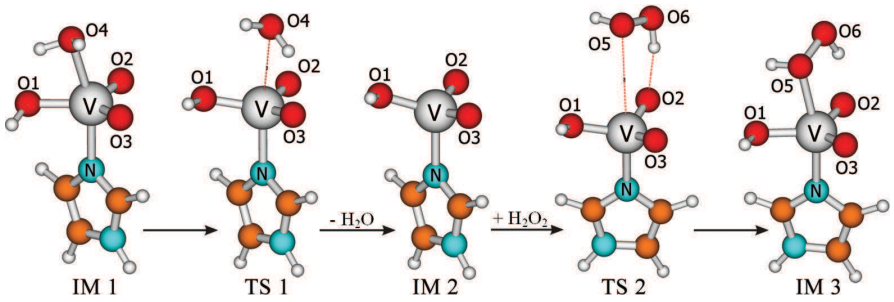


Figure 6. Intermediate and transition state structures for the reaction of hydrogen peroxide with imidazole-bound vanadate(V) optimized at ADF/BP86/TZ2P level. Analogous calculations have been done on imidazole bounded molybdate(VI).

larger coordination numbers compared to vanadium, it can be assumed that it prefers an associative reaction mechanism while vanadium prefers a dissociative one. In order to verify this claim, structures of several stationary points on the potential surface along the reaction coordinate for this first dissociative step of the catalytic cycle were calculated for both the vanadium and the molybdenum species (Fig. 6). Activation energy barriers and Gibbs free reaction energy values for both species were calculated from the free energy values of the optimized structures. Indeed an important difference is obtained: The reaction of hydrogen peroxide with $[VO_2(OH)(H_2O)Im]$ is an *exergonic* process while the same reaction employing $[MoO_2(OH)(H_2O)Im]^+$ as substrate is *endergonic*.

$$\textbf{IM1}(\mathrm{V}) + \mathrm{H_2O_2} \rightleftharpoons \textbf{IM3}(\mathrm{V}) + \mathrm{H_2O} \qquad \triangle_{\mathrm{R}}\mathrm{G} = -18,6\ kJ/mol \qquad (2)$$

$$\textbf{IM1}(\mathrm{Mo}) + \mathrm{H_2O_2} \rightleftharpoons \textbf{IM3}(\mathrm{Mo}) + \mathrm{H_2O} \qquad \triangle_{\mathrm{R}}\mathrm{G} = +24,0\ kJ/mol \qquad (3)$$

Without the formation of a peroxo complex, the catalytic reaction at the metal centre cannot proceed. This is the reason why molybdate does not show catalytic activity when embedded into the VHPO apoenzyme

6 Future Prospects

Investigations of the enzymatic cycle of vanadium haloperoxidases (Fig. 4) have so far been conducted mainly by calculations employing the small model system (Fig. 3 b). Expanding the system by the surrounding enzyme matrix (Fig. 3 d), the important hydrogen bonding network between the cofactor and several amino acid residues can be properly taken into account. Some previously obtained results can be verified using this larger model system. The TURBOMOLE suite of programs[18], now available for use on the NIC cluster JUMP, is especially well suited for these calculations on large systems as it contains the *multipole acceleration for the resolution of identity* method (MARIJ).

Calculations merging quantum chemical methods and molecular mechanics (QM/MM) represent one further step to improve the model system. MM calculations require only a fraction of the computational time needed for QM calculations. Using this approach, it is possible to model the full enzyme matrix (with MM) instead of only a small section at the active site, and thus no atoms of the backbone need to be fixed. The surrounding area of the active site is still treated in a fully quantum chemical way.

Additionally, it is desirable to take a closer look at the associative mechanism of both molybdate(VI) and vanadate(V) species, thus gaining knowledge to specifically improve ligands in order to produce molybdenum(VI) complexes exhibiting strong haloperoxidase activity.

Acknowledgments

This project is supported by a grant of computer time of the John von Neumann Institute for Computing (NIC) and a research grant of the Deutsche Forschungsgemeinschaft.

References

1. A. Messerschmidt and R. Wever, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 392–396.
2. W. Plass, *Angew. Chem. Int. Ed.* **1999**, *38*, 909–912.
3. A. Messerschmidt, L. Prade, and R. Wever, *Biol. Chem.* **1997**, *378*, 309–315.
4. W. Plass, A. Pohlmann, and H.-P. Yozgatli, *J. Inorg. Biochem.* **2000**, *80*, 181–183.
5. M. Bangesh and W. Plass, *J. Mol. Struct. THEOCHEM* **2005**, *725*, 163–175.
6. G. Zampella, P. Fantucci, V. L. Pecoraro, and L. De Gioia, *J. Am. Chem. Soc.* **2005**, *127*, 953–960.
7. S. Macedo-Ribeiro, W. Hemrika, R. Renirie, R. Wever, and A. Messerschmidt, *J. Biol. Inorg. Chem.* **1999**, *4*, 209–219.
8. G. J. Colpas, B. J. Hamstra, J. W. Kampf, and V. L. Pecoraro, *J. Am. Chem. Soc.* **1996**, *118*, 3469–3478.
9. M. J. Clague and A. Butler, *J. Am. Chem. Soc.* **1995**, *117*, 3415–3404.
10. G. E. Meister and A. Butler, *Inorg. Chem.* **1994**, *33*, 3269–3275.
11. A. F. Ghiron and R. C. Thompson, *Inorg. Chem.* **1990**, *29*, 4457–4461.
12. J. D. Lydon, L. M. Schwane, and R. C. Thompson, *Inorg. Chem.* **1987**, *26*, 2606–2612.
13. A. Butler and A. H. Baldwin, Structure and Bonding – Metal Sites in Proteins and Models; Springer, **1997**, pp. 111–132.
14. H. Vilter, *Phytochemistry* **1984**, *23*, 1387–1390.
15. ADF Version 2004-07, Scientific Computing & Modelling, Theoretical Chemistry, Virje University, Amsterdam, `http://www.scm.com`.
16. C. F. Guerra, J. Snijders, G. te Velde, and E. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391–403.
17. G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonesca, C. F. Guerra, E. J. Baerends, J. G. Snijders, and T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931–967.
18. R. Ahlrichs, M. Bar, M. Haser, H. Horn, and C. Kolmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
19. A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
20. J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822–8824.
21. E. van Lenthe, E.-J. Baerends, and J. G. Snijders, *J. Chem. Phys.* **1994**, *101*, 9783–9792.
22. D. Andrae, U. Haeussermann, M. Dolg, H. Stoll, and H. Preuss, *Theor. Chim. Acta* **1990**, *77*, 123–141.
23. W. Plass, M. Bangesh, S. Nica, and A. Bucholz, *ACS Symposium Series* **2007**, *974*, 163–177.

Chemistry: Facing Complexity

Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich
USI Campus, Via Giuseppe Buffi 13
CH-6900 Lugano, Switzerland
E-mail: parrinello@phys.chem.ethz.ch

The fundamental laws that regulate the chemical-physical behaviour of condensed systems under ordinary conditions of temperature and pressure are well known. However, it is only in the last few decades that appropriate algorithms have been devised to allow the realistic simulation of complex systems based on first principles simulations. This has led to the development of efficient codes capable of exploiting the power of modern parallel machines. In the following four reports a number of groups present the impressive results of their work, obtained using the NIC machines. If we have to find a unifying feature in these varied papers, it is the effort to treat complex problems. The complex systems studied are modelled in as realistic a manner as possible. This is permitted by the state-of-the-art resources available and the development of clever algorithms. In this respect the use of advanced sampling methods such as metadynamics is noteworthy. The net effect of these technical innovations is very impressive; suffice it here to look at the unprecedented scope of the work of Professor Marx' group. I believe that these papers give a taste of things to come and of the brave new world where ever more complex and difficult problems can be tackled with a combination of computer power and clever algorithms.

A Density-Functional Study of Nitrogen and Oxygen Mobility in Fluorite-Type Tantalum Oxynitrides

Holger Wolff, Bernhard Eck, and Richard Dronskowski

Institut für Anorganische Chemie, RWTH Aachen
Landoltweg 1, 52056 Aachen, Germany
E-mail: {H. Wolff, B. Eck, R. Dronskowski}@ac.rwth-aachen.de

In this contribution we present results of our theoretical studies of nitrogen mobility in solid M–Ta–O–N systems. Periodic supercell calculations at density-functional level have been performed to investigate the local structure of N-doped oxynitrides and the anion diffusion mechanisms. The migration pathways and activation barriers were calculated using the nudged elastic band method with the climbing-image enhancement. We show that the defect migration is mainly caused by the diffusion of oxygen anions. The activation energy can be lowered by increasing the defect concentration, and it is, to a large extent, depending on the dopant size.

1 Introduction

Solid electrolytes are desirable in cells or batteries where liquid electrolytes may be undesirable, such as in implantable medical devices. Preferred solid electrolytes include materials that are solid at room temperature, electrically insulative and ionically conductive. The ionic conductivity is strongly affected by the anion vacancy concentration which is enhanced by doping with aliovalent cations or with nitrogen[1,2]. Anion diffusion in ternary Zr–O–N and quaternary Y–Zr–O–N materials has been investigated experimentally by single crystal neutron diffraction[3], tracer diffusion[4,5] and impedance spectroscopy[2]. Only a small number of theoretical studies has been performed so far on M–O–N systems, mainly focused on the description of well-defined bulk structures. It is well-known that lattice distortions determine the ionic conductivity to a large extent[6,7], and it has been argued that conductivity activation energies depend on vacancy ordering in the anion sublattice[8].

Recently, yttrium-doped tantalum oxynitrides have been synthesized for the first time, and these materials crystallize with defect fluorite-type structures[9,10]. The energy barriers for ionic conductivity in such compounds have been calculated from first principles, and we present a comparative view touching upon dopant size, the amount of dopants, lattice distortions and vacancy ordering.

2 Computational Method

2.1 Modelling

A twofold unit cell containing eight formula units of $M_x\mathrm{Ta}_{1-x}(\mathrm{O,N,}\square)_2$ was used in this study. The migration pathway of the vacancy and, also, the energy barrier to migration were determined by finding the minimum energy path from one lattice site to an adjacent site using the nudged elastic-band method[6] in its climbing image variant[7,11,12] as implemented in the density-functional Vienna *Ab initio* simulation package (VASP)[13–16].

First, the starting and final configurations were determined by optimizing the structure with the vacancy at each of the two adjacent sites. A set of seven intermediate configurations were then optimized such as to converge to points lying on the minimum-energy path. All nine configurations are shown in Figure 1. Harmonic-spring interactions were used to connect the different images of the system along the path. Calculations were performed using the GGA functional of Perdew and Wang[17] and ultrasoft pseudopotentials of Vanderbilt type[18]. The kinetic cutoff-radius was chosen to be 490 eV, calculating Gamma point only with about 70,000 plane-waves in a total of about 70 bands included, the exact numbers depending on the system beheld.

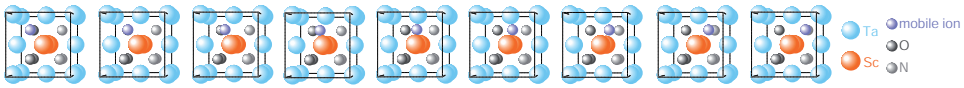


Figure 1. Initial guess for the calculation of a diffusion path in $M_{0.125}\mathrm{Ta}_{0.875}\mathrm{O}_{0.875}\mathrm{N}\square_{0.125}$

2.2 System Requirements

To perform the nudged elastic band method with VASP, the number of CPUs must be dividable by the number of intermediate configuration images. VASP divides the processors into groups with each group calculating one image. Here, the total number of iterations until convergence is reached can be up to 50,000.

Earlier performance tests on a Sun Fire E25K-Cluster concerning the parallization performance of VASP have shown that an efficiency of $t/t_{\mathrm{opt}} = 99\%$ can be reached up to 16 CPUs per image, so that a maximum of $7 \times 16 = 112$ CPUs makes economic sense. On the IBM p690 Cluster JUMP, it is sensible to calculate each image on one complete node and thus minimizing ethernet traffic, giving a request for $7 \times 4 = 28$ CPUs. A performance test with a total of only 7 processors has shown that, on the IBM p690 Cluster, calculations undergo a speedup of 40 % compared with a Sun Fire E25K-Cluster.

Concerning memory requirement, VASP relies on dynamic memory allocation. From experience, the charge density symmetrization takes usually the highest amount of memory. Additionally, the executable and all tables have to be held on all processors as well as a complex array of the size $N_{\mathrm{bands}} \times N_{\mathrm{bands}}$ leading to an average memory allocation of about 4 GBytes per CPU and peak memory allocations of up to 9 GBytes per CPU.

3 Results and Discussion

In the system Ta–O–N, the fluorite-type structure is stabilized upon doping with medium-sized trivalent cations such as rare-earth elements[19]. For reasons of electroneutrality, a number of vacancies is introduced into the anionic sublattice, making these compounds promising candidates for ionic (nitrogen and oxygen) conductors. It is often argued that lattice distortions due to ionic radii mismatch lead to an increasing activation barrier for ionic conductivity[20]. In order to study this influence of dopant size for the case of fluorite-type

M	r_{eff} (Å)
Fe	0.78
Sc	0.87
In	0.92
Y	1.02
Ta	0.74

Table 1. Effective ionic radii as taken from reference[21].

tantalum oxynitrides, we modelled several – partly hypothetical – fluorite-related compounds in the system M–Ta–O–N with four different cations M, namely Fe^{3+}, Sc^{3+}, In^{3+}, and Y^{3+}. The ionic radii[21] of Fe^{3+} and Ta^{5+} are almost identical and they are slightly increasing towards Y^{3+} (see Table 3).

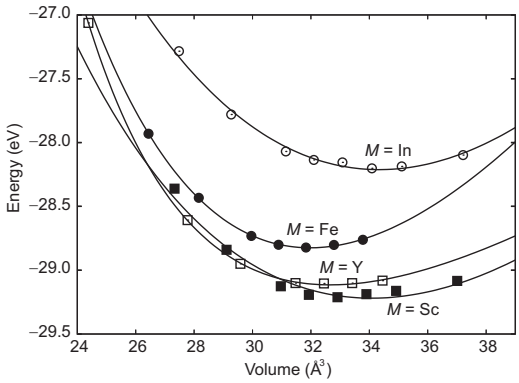


Figure 2. Total Energies for differently doped tantalum oxynitrides $M_{0.125}\text{Ta}_{0.875}\text{O}_{0.875}\text{N}\square_{0.125}$

First, structural relaxation was allowed to ensure that the chosen compounds make sense in terms of structure and stability. The resulting structures were stable but heavily distorted; nonetheless, the classification of these compounds as being of a defect flourite-type is sustainable[10]. The energy–volume functions are shown in Figure 2 for the example of $M_{0.125}\text{Ta}_{0.875}\text{O}_{0.875}\text{N}\square_{0.125}$.

The migration pathways and activation barriers were then calculated for both cases of oxygen and nitrogen movement. In the fluorite structure, the anions are located in the tetrahedral sites within the face-centred cubic cation framework. Judging from chemical intuition, one would expect that the minimum-energy path for the diffusion of an anion to an adjacent vacant site touches the empty octahedral sites, resulting in *two* saddle points on the energy hypersurface according to the two sharing faces of the octahedron and the two tetrahedra. This is very well reproduced by the DFT calculations (compare Figure 3).

Comparing the activation barriers for oxygen and nitrogen movement (see Table 3), it turns out that oxygen is generally more mobile, which can be traced back to the higher

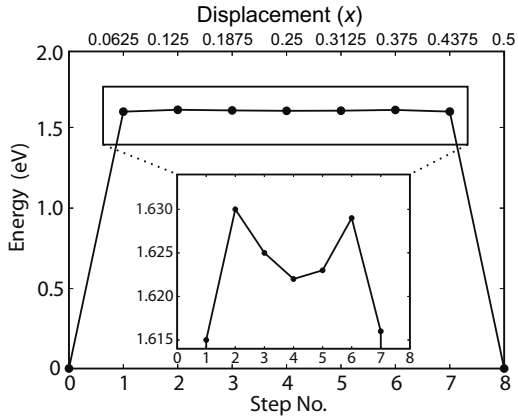


Figure 3. Calculated energies along the minimum-energy path for oxygen migration in $Sc_{0.25}Ta_{0.75}O_{0.75}N\square_{0.25}$. The central parts are enlarged in the middle.

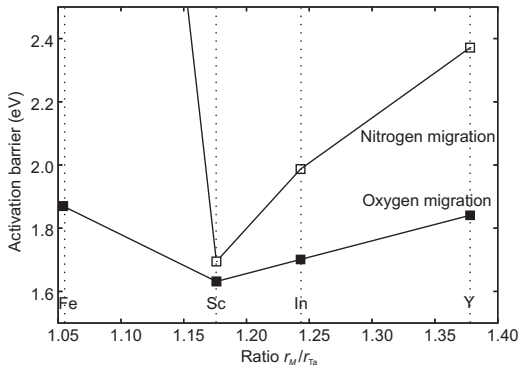


Figure 4. Calculated energy barriers for nitrogen migration (empty boxes) and oxygen migration (filled boxes) in $M_{0.25}Ta_{0.75}O_{0.75}N\square_{0.25}$, as a function of the ratio of the ionic radii r_M/r_{Ta}.

charge of the N^{3-} anion. Decreasing the dopant size and thus approximating a ratio $r_M/r_{Ta} \approx 1$, the energy threshold lowers, as can be seen in Figure 4. While this is true for Sc, In and Y, the anionic movement in the Fe-doped compounds is dramatically hindered. Compared with the other three dopant cations, which achieve noble gas configuration in their trivalent state, Fe^{3+} has $3d^5$ configuration. Thus, the lowered anionic mobility probably goes back to the more covalent bonding between Fe and N.

Increasing the dopant amount leads to a growing number of vacancies, which causes a higher ion mobility and a decrease of the activation barrier. In Figure 5 the energy barriers are shown in relation to both dopant amount and vacancy concentration. It is obvious that the energy threshold is independent of the dopant concentration but, instead, determined by the vacancy concentration: Turning from $M_{0.25}Ta_{0.75}O_{0.75}N\square_{0.25}$ to $M_{0.375}Ta_{0.625}ON_{0.75}\square_{0.25}$ the activation energy does not significantly change, since

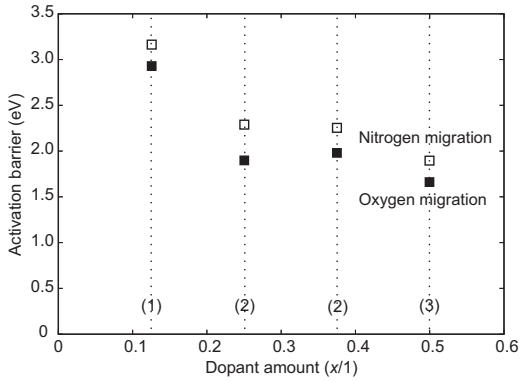


Figure 5. Energy barriers for nitrogen migration (empty boxes) and oxygen migration (filled boxes) in $Y_xTa_{1-x}(O,N,\square)_2$, as a function of the dopant amount and the number of vacancies per unit cell (in parentheses).

the number of vacancies is two in both unit cells. In $M_{0.5}Ta_{0.5}O_{0.875}N_{0.75}\square_{0.375}$, on the other hand, one additional defect is brought into the unit cell, and the barrier is again lowered.

4 Conclusions

Density-functional calculations show that the energy barrier for anion diffusion in fluorite-type tantalum oxynitrides is closely related to both dopant size and dopant amount. Increasing the dopant amount causes a higher vacancy concentration, making it easier for the ions to move inside the material. The ratio of the cationic radii determines the activation threshold to a very large extent, making it decrease while r_M/r_{Ta} approaches unity. The defect migration, however, is mainly caused by the diffusion of oxide anions, which are slightly more mobile than the nitride anions.

Acknowledgments

The authors wish to thank the computing centres at RWTH Aachen University, Jülich Research Centre and HLRS Stuttgart for providing large amounts of CPU time. This project is supported by Deutsche Forschungsgemeinschaft in the priority program “Substitutional effects in ionic solid”.

References

1. M. Lerch, *Nitridation of Zirconia*, J. Am. Ceram. Soc. **79**, 2641-2644, 1996.
2. J.-S. Lee, M. Lerch, J. Maier, *Nitrogen-doped Zirconia: A comparison with cation stabilized zirconia*, J. Solid State Chem. **179**, 270-277, 2006.
3. I. Kaiser-Bischoff, H. Boysen, C. Scherf, T. Hansen, *Anion diffusion in Y- and N-doped ZrO_2*, Phys. Chem. Chem. Phys. **7**, 2061-2067, 2005.

4. M. Kilo, C. Argirusis, G. Borchardt, R.A. Jackson, *Oxygen diffusion in yttria stabilised zirconia—experimental results and molecular dynamics calculations*, Phys. Chem. Chem. Phys. **5**, 2219-2224, 2003.
5. M. Kilo, M.A. Taylor, C. Argirusis, G. Borchardt, M. Lerch, O. Kaitasov, B. Lesage, *Nitrogen diffusion in nitrogen-doped yttria stabilised zirconia*, Phys. Chem. Chem. Phys. **6**, 3645-3649, 2004.
6. G. Mills, H. Jónsson, G.K. Schenter, *Reversible work transition state theory: Application to dissociative adsorption of hydrogen*, Surf. Sci **324**, 305-337, 1995.
7. G. Henkelman, H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points*, J. Chem. Phys. **113**, 9978-9985, 2000.
8. E.D. Wachsmann, *Effect of oxygen sublattice order on conductivity in highly defective fluorite oxides*, J. Europ. Ceram. Soc. **24**, 1281-1285, 2004.
9. H. Schilling, H. Wolff, R. Dronskowski, M. Lerch, *Fluorite-type Solid Solutions in the System Y–Ta–O–N: A Nitrogen-rich Analogue to Yttria-stabilized Zirconia (YSZ)*, Z. Naturforsch. B **61**, 660-664, 2006.
10. H. Wolff, H. Schilling, M. Lerch, R. Dronskowski, *A Density-Functional and Molecular Dynamics Study on the Properties of Yttrium-doped Tantalum Oxynitride*, J. Solid State Chem. **179**, 2265-2270, 2006.
11. G. Henkelman, B.P. Uberuaga, H. Jónsson, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*, J. Chem. Phys. **113**, 9901-9904, 2000.
12. H. Jónsson, G. Mills, K.W. Jacobsen, *Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions* in Classical and Quantum Dynamics in Condensed Phase Simulations, B.J. Berne, G. Cicotti, D.F. Coker (Eds.), World Scientific, 1998.
13. G. Kresse, J. Hafner, *Ab initio molecular dynamics for liquid metals*, Phys. Rev. B **47**, 558-561, 1993.
14. G. Kresse, J. Hafner, *Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium*, Phys. Rev. B **49**, 14251-14269, 1994.
15. G. Kresse, J. Furthmüller, *Efficiency of ab initio total energy calculations for metals and semiconductors using a plane-wave basis set*, Comput. Mat. Sci. **6**, 15-50, 1996.
16. G. Kresse, J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B **54**, 11169-11186, 1996.
17. J.P. Perdew, Y. Wang, *Accurate and simple analytic representation of the electron-gas correlation energy*, Phys. Rev. B **45**, 13244-13249, 1992.
18. D. Vanderbilt, *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism*, Phys. Rev. B **41**, 7892-7895, 1990.
19. P. Maillard, F. Tessier, E. Orhan, F. Cheviré, R. Marchand, *Thermal Ammonolysis Study of the Rare-Earth Tantalates $RTaO_4$*, Chem. Mater. **17**, 152-156, 2005.
20. M. Mogensen, D. Lybye, N. Bonanos, P.V. Hendriksen, F.W. Poulsen, *Factors controlling the oxide ion conductivity of fluorite and perovskite structured oxides*, Solid State Ionics **174**, 279-286, 2004.
21. R.D. Shannon, *Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalkogenides*, Acta Cryst. A **51**, 751-767, 1976.

First-Principles Molecular Dynamics Simulations of Liquids and Solvent Effects

Jens Thar and Barbara Kirchner

Wilhelm-Ostwald Institut für Physikalische und Theoretische Chemie
Universität Leipzig, Linnestr. 2, D-04103 Leipzig
E-mail: bkirchner@uni-leipzig.de

We present several *first-principles* molecular dynamics simulations studies of liquids and solvent effects. The first project deals with dipole moments of one alanine molecule dissolved in water molecules. The average dipole moment of the water molecules depends on their vicinity to the distinct functional groups of alanine. The results of the *first-principles* molecular dynamics simulation is used to verify a cluster ansatz. In this ansatz, snapshots of alanine surrounded by different shells of water molecules are cut out of a trajectory. Sufficient convergence of the dipole moment of alanine and its first water shell can be found for two solvent shells. In the second project, we examine the basis set superposition error along a free-energy trajectory of the water dissociation reaction. Snapshots are again cut out of a trajectory and their basis set superposition error is calculated by static quantum mechanics. We observe an increase of the basis set superposition error with the decline of the distance of the two molecules. In the third project which is still in progress, we investigate an example of the third generation ionic liquid, ethyl-methyl-imidazoliumthiocyanate ($[C_2C_1im]SCN$).

1 Introduction

1.1 Alanine in Water[1]

The solvation behaviour of amino acids in water is one of the key questions to be addressed in biology, due to its drastic influence on their chemical reactivity. The most straightforward way to adequately describe such a flexible system are *first-principles* molecular dynamics simulations. In this project the electronic structure is accurately calculated using density functional theory, while the configurational space is explored by an efficient and temperature dependent sampling technique.

Another possibility to derive properties of molecules dissolved in liquids is the so called cluster ansatz where a trajectory either obtained by classical or first-principles molecular dynamics simulations samples the configurations in the solvents. This trajectory is subsequently harvested for small clusters. The main concern with this approach is the number of snapshots and the number of solvent molecules needed to achieve convergence with respect to the experiment.

1.2 BSSE Error along the Free Energy Trajectory[2]

Atom-centreed basis sets (e.g. the Gaussian Basis Set (GBS)) that are used in static quantum calculations introduce the so-called basis set superposition error (BSSE), when comparing interaction energies (E_I) at different geometries. The effect of this error on the static geometries is under control, because it remains nearly constant close to the minimum structure. In simulations, the average structure is determined by sampling a broader

region around the local minima and by weighting configurations according to their energy. Therefore, the BSSE is not necessarily constant, and might have a direct influence on the weighting and hence the average structure.

1.3 Simulation of Third Generation Ionic Liquids

The wide applicability of ionic liquids as environmentally friendly solvents has led to a steep increase of publications on this subject. Ionic liquids are fused salts of organic cations and most often inorganic anions.[3,4] They are good solvents for a wide range of inorganic and organic reactions. The reason for this is attributed to their microscopic properties. These properties are mainly their high polarity and their non-coordinating behaviour.

2 Scope

2.1 Alanine in Water

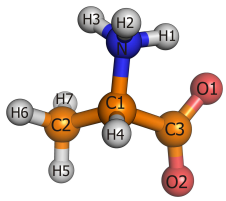


Figure 1. Structure of alanine with atom labels as used in this chapter.

In this project, we investigate alanine as the smallest biologically active α-amino acid. The zwitterionic form of the alanine shows three sites with distinct polarity: The basic COO^- group, the acidic NH_3^+ group, and the aliphatic CH_3 group. The main focus is the development of the dipole moment of a zwitterionic alanine in water as well as of different water shells along a trajectory. These local dipole moments are estimated by *first-principles* simulation in combination with maximally localized Wannier centres (MLWC). Additionally, we want to test to what extent the cluster ansatz not relying on periodic boundary conditions and with fewer solvent molecules produces comparable results. Therefore, we cut out snapshots with different numbers of solvent molecules along the simulated trajectory and calculate local dipole moments by means of MLWCs. This allows us to verify the cluster ansatz against the results from our previous *first principles* simulation.

2.2 BSSE Error along the Free Energy Trajectory

We investigate the reaction

$$\mathrm{H_2O^\star \cdots H^\star\text{–}O\text{–}H \rightarrow H_3O^+ + OH^-}$$

by means of thermodynamical integration. Despite the fact that this reaction does not take place in the gas phase, it serves us to obtain many chemically different situations. During its course, each of these situations would be associated with a unique BSSE when a Gaussian type orbital basis set would be applied. The trajectory is therefore harvested for these different configurations, and their BSSE is determined by means of static quantum mechanics.

2.3 Simulation of Third Generation Ionic Liquids

Here, we investigate a third generation ionic liquid, ethyl-methylimidazoliumthiocyanate ([C_2C_1im]SCN).

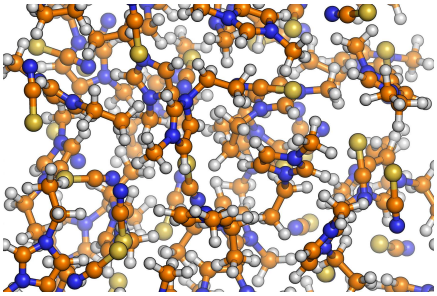


Figure 2. Structure of ethyl-methylimidazoliumthiocyanate.

A *first-principles* molecular dynamics simulation can determine the main building blocks of the liquid and their life time as well as the diffusion constant of the dominating species.

3 Methodology

We simulate one alanine as well as the free energy trajectory with the CPMD code, which is based on the Car–Parrinello method [6]. The electronic structure is described in the framework of density functional theory, utilising the gradient corrected exchange correlation functional BP86. The Kohn–Sham orbitals are expanded in a plane wave (PW) basis with a kinetic energy cutoff of 70 Rydberg, while core electrons are approximated by norm conserving pseudo potentials of the Troullier–Martins type[5]. Local dipole moments for CPMD calculations are derived based on maximally localised Wannier centres (MLWC)s.

For alanine, the system consists of 60 water molecules surrounding the alanine in a periodic box. The equilibration simulation was thermostated at 320 K using Nosé–Hoover thermostats[7–9], while the production run was performed in the NVE ensemble. Average computer time for a Car–Parrinello molecular dynamics step on 16 CPUs of the JUMP Cluster was 7.4 seconds. Additional averaged time needed for calculation of the maximally localized Wannier centres (MLWC)s was 14.8 seconds.

Concerning the BSSE, this error is estimated by the counterpoise correction (CP) of the dissociation reaction given in Eq. (2.2). Each monomer m is calculated in the basis set

centred at its atoms (E_m^{mon}) and in the complete basis set (E_m^{all}). The CP of the reaction is defined by the difference $(E_1^{mon} + E_2^{mon}) - (E_1^{all} + E_2^{all})$.

For [C_2C_1im]SCN, molecular dynamics simulation using the GPW[10] ansatz in the framework of the CP2K/Quickstep[11] environment has been performed. Temperature was set to 398.15 K and controlled by a Nosé–Hoover thermosthat. We use GTH type pseudopotentials[12,13] to describe the core electrons of all atoms. A DZVP basis set is used for the representation of the Kohn–Sham orbitals, and upon switching to the plane wave representation, a cutoff of 300 Ryd is used. We use the gradient corrected functional PBE throughout the simulation.

4 Results: Alanin

4.1 Dipole Moments from Molecular Dynamics: Water

We investigate the dipole moment of water molecules around each of the three distinct functional groups NH_3^+, CH_3 and COO^-. The probability distributions of the dipole moment from different shells inside each of these cones is depicted in Fig. 3. The first shell around the oxygen atom contains all water molecules up to a distance of 330 pm, the second shell contains all water molecules from 330 pm up to a distance of 540 pm, and further solvent molecules are collected in the third shell. For nitrogen, these two cutoffs are set to 360 pm and 590 pm, and for the carbon atoms, the first shell consists of water molecules up to a distance of 510 pm, and all further solvent molecules are collected in the second shell.

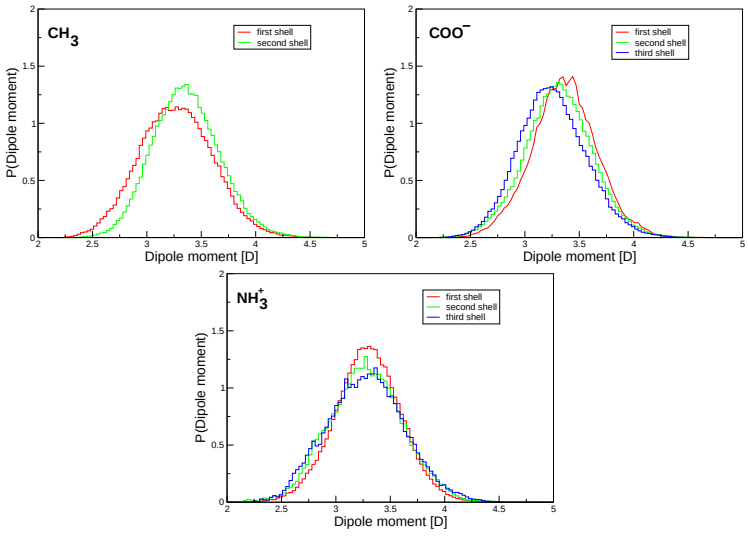


Figure 3. Probability distribution function of the dipole moments of water molecules closest to atoms of the COO^-, the CH_3 and the $NH3^+$ group of the alanine. See text for details.

As expected, the average dipole moment of the solvent molecules in the vincity of the

COO^- decays from 3.35 D (first shell) over 3.31 D (second shell) to 3.25 D (third shell) with rising distance from the solute. In contrast, the average dipole moments of the solvent molecules around the CH_3 and the NH_3 group of alanine rise from 3.27 D (3.28 D) to 3.36 D (3.0 D) when entering the second shell.

4.2 Dipole Moments from Static Calculations: The Cluster-Ansatz

In this project we want to explore whether the use of periodic boundary conditions is necessary for the computation of electronic properties of the solute (i.e. the alanine) or if these properties are sufficiently converged already with two or even less layers of solvent molecules added. i.e., we want to investigate for which cluster size the cluster method is valid.

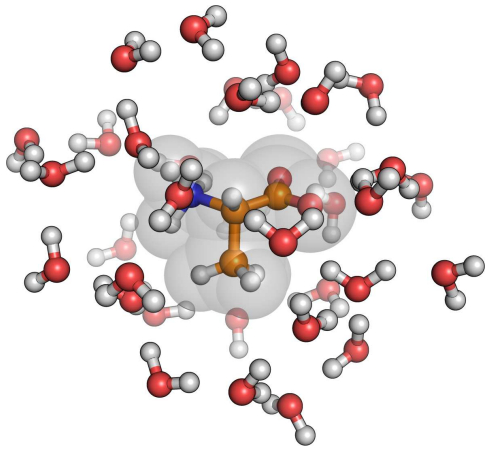


Figure 4. Snapshot of one set solvent molecules.

Therefore, we selected 4 distinct sets of snapshots: The first two sets only account for the first (and second) solvent shell of the two polar groups. The second two sets account for the first (and second) solvent shell of all selected subgroups. A short assay revealed that already 50 snapshots per set are sufficient to converge the average dipole moments.

Method	**set0**	**set1**	**set2**	**set3**	**set4**	**full trajectory**
PW/MLWC	10.99	14.10	15.08	14.59	15.56	15.93

Table 1. Average local dipole moments in D of alanine within different sets of solvent molecules.

The influence of different solvent shells on the local dipole moment of alanine is summarized in table 1. The dipole moment converges nicely with increasing number of electrons. As a result, **set4** is already a reasonable good approximation to the periodic boundary conditions, yielding about 98% of the reference dipole moment. A similar analysis

has been performed for the dipole moment of the water molecules around the first shell of the alanine. We obtain a dipole moment of 3.15 D, about 95% of the reference simulation's value, for the innermost water molecules in **set4**. Thus, the cluster ansatz provides a reliable tool to assess properties of systems with a large number of degrees of freedom.

5 Results: BSSE Error along the Free Energy Trajectory

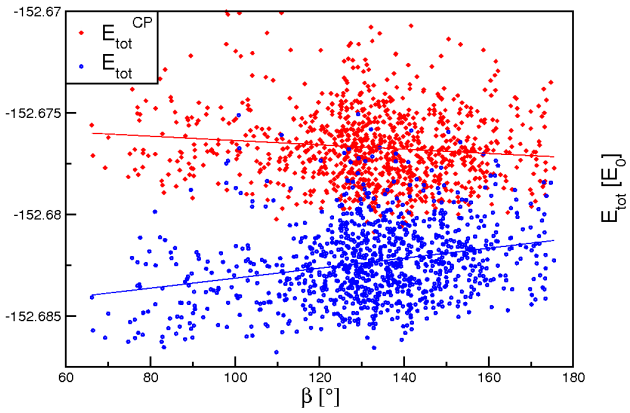


Figure 5. Total energies along the constraint trajectory r1=180 pm plotted against the angle β E_{tot}: uncorrected total SVP energy; E_{tot}^{CP}: counterpoise corrected total SVP energy.

In Fig. 5 we plotted the total energies as well as the total energies with subtracted counterpoise corrections against the angle β for the part of the free energy simulation where the distance between the two products is restrained to 180 pm. β denotes the angle between the vector pointing from the oxygen atom of the first to the oxygen atom of the second water molecule and the vector pointing from the oxygen of the second water molecule to its centre of mass. The corrected values (red) are higher in the graph than the uncorrected values (blue). The straight lines in Fig. 5 show linear regressions in order to clarify the tendencies. It is apparent from the Figure 5 that configurations with small distances between atoms of different molecules (represented here by small angles β) would be overpopulated since their energy (blue) is more favourable than if correct energies (red) would be calculated, i.e., the BSSE is not a simple shift of the potential energy surface.

6 Results: Structure of Bulk ([C_2C_1im]SCN)

In Fig. 6 the radial pair distribution function g(r) of different atoms of the cation ethyl-methyl-imidazolium is given. Namely, the distances between the nitrogen and sulfur of the thiocyanate and the imidazolium ring's three protons of the cation are plotted. The exact definition of each atom is displayed in Fig. 2. We see that the first peak at about 220 pm, representing the smallest distance between the two ions, is part of the $H_2 - N$ g-function.

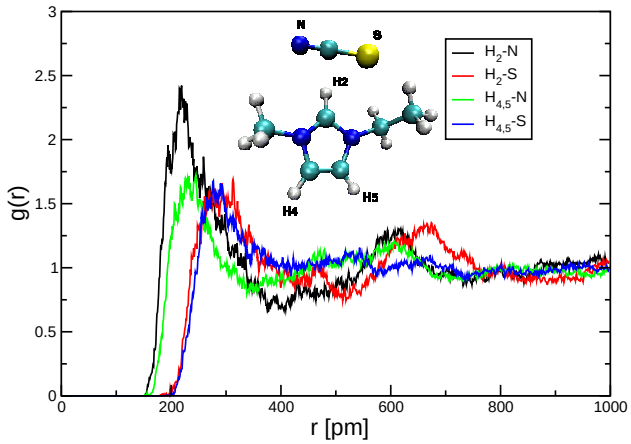


Figure 6. Selected g-function of $[C_2C_1im]SCN$.

The second peak at about 230 pm is provided by the $H_{4,5} - N$ radial pair distribution function. Therefore, we conclude that the preferred coordination site of the cation is the H_2 and of the thiocyanate it is the nitrogen atom. The first "shell" of sulfur atoms around the cation is found at distances of about 280 pm ($H_{4,5} - S$) and 290 pm ($H_2 - S$). It is interesting that this order is opposed to the order found for the nitrogen atoms binding to H_2 or $H_{4,5}$. We must note that according to the average over the autocorrelation function for the orientation of all SCN^- anions' S-N vectors, the simulation is still not converged, and the data presented is still preliminary.

7 Conclusion

We have presented the results of various projects in the field of *first principles* molecular dynamics simulations. Concerning alanine in water, we found that the large dipole moment of alanine induces a considerable increase of the dipole moments of the surrounding water molecules. Also, the different functional groups of alanine exert different but clearly observable influence on their surrounding water molecules. Our test with the cluster ansatz showed that at least 40 water molecules have to be included to converge properties like the dipole moment. Our second project, the impact of the BSSE on simulations using Gaussian basis sets, showed that a sizable BSSE should result in an overestimation of structures which are relatively close together. The third project still runs on JUMP and additional data might be available at the end of this accounting period.

8 Performance on NIC

As a measurement of the performance of the NIC cluster for CP2K, we plotted the average time step of one of our simulations of $[C_2C_1im]SCN$ multiplied with the number of used CPUs against this number, see Fig. (7). The graph shows that CP2K scales favorably for up to 64 CPUs. Using one 16 CPUs is the most time efficient case.

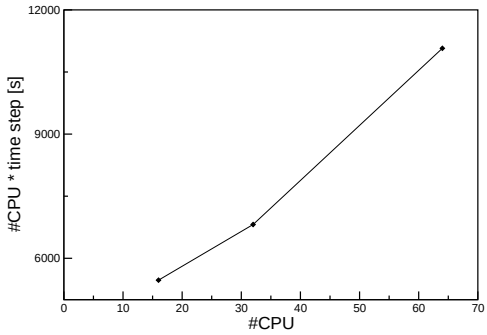


Figure 7. Number of CPUs times the average time step for a CP2K run ploted against the number of CPUs.

Acknowledgments

The authors gratefully acknowledge the financial support of the DFG priority program SPP 1191 “Ionic Liquids” and the ERA Chemistry program, that allows fruitful collaboration under the project “A Modular Approach to Multi-responsive Surfactant/Peptide (SP) and Surfactant/Peptide/Nanoparticle (SPN) Hybrid Materials”. We furthermore like to thank the collaborative research center SFB 624 “Templates” at the University of Bonn for financial support. This work was supported with computational time provided by the NIC supercomputers in Jülich. J.T. thanks funding by a “Chemiefonds-Stipendium” of the Fonds der Chemischen Industrie.

References

1. Thar, J.; Zahn, S.; Kirchner, B. *J. Chem. Phys. B, in print*.
2. Thar, J.; Hovorka, R.; Kirchner, B. *J. Chem. Theor. Comp.* **2007,** *3,* 1510.
3. T. Welton. *Chem. Rev.* **99** (1999) 2071.
4. T. Welton. *Coord. Chem. Rev.* **248** (2004) 2459.
5. Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991,** *43,* 1993.
6. Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985,** *55,* 2471.
7. Nosé, S. *J. Chem. Phys.* **1984,** *81,* 511.
8. Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992,** *97,* 2635.
9. Hoover, W. G. *Phys. Rev. A* **1985,** *31,* 1695 - 1697.
10. Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **103** (1999) 124.
11. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005,** *167,* 103.
12. Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **58** (1998) 3641.
13. Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **54** (1996) 1703.

First-Principles Lattice Dynamics Calculations of the Phase Boundary Between β-Si_3N_4 and γ-Si_3N_4 at Elevated Temperatures and Pressures

Atsushi Togo and Peter Kroll

Institut für Anorganische Chemie
Rheinisch-Westfälische Technische Hochschule Aachen (RWTH)
Landoltweg 1, 52056 Aachen, Germany
E-mail: peter.kroll@ac.rwth-aachen.de

The phase boundary between β-Si_3N_4 and γ-Si_3N_4 is investigated at high pressure and high temperature using first-principles lattice dynamics calculations within the quasi-harmonic approximation. We find a positive slope of the phase boundary. It turns out that the thermal expansion of the spinel-type γ-phase is larger than that of the phenacite-type β-phase. On the other side, pressure affects more the volume of β-Si_3N_4 than of γ-Si_3N_4, reflected in the higher bulk modulus of γ-Si_3N_4. The origin of the different temperature behaviour of these phases, consequently, goes along with a larger volume dependence of the zero point energy in γ-Si_3N_4 in comparison to β-Si_3N_4.

1 Introduction

Almost a decade ago cubic γ-Si_3N_4 with spinel-type structure was realized by high-temperature/high-pressure chemical synthesis[1]. The first report gave a pressure of 15 GPa and a temperature of 2000 K needed to favour the formation of the γ-phase over the β-phase. Computations that went along with the experimental realization indicated a transition pressure of 12.1 GPa. Subsequent experimental studies, which used amorphous silicon and nitrogen containing precursors to exclude kinetic effects of the β-γ transition, located the phase boundary between 10 and 13 GPa at temperatures between 1600 and 2000 K[2]. No account for a direct conversion from the γ-phase to the β-phase has been given, though.

At high temperatures and pressures, it is usually extremely difficult to determine correctly the phase boundaries of solids because large hysteresis effects due to sluggish kinetics are usually present. In this study, we compute the pressure-temperature phase diagram between β- and γ-Si_3N_4 assuming ideal crystals using first-principles lattice dynamics calculations. The crystal structures of β- and γ-Si_3N_4 are shown in Figures 1 and 2, respectively. Si atoms in the phenacite-type β-structure are four-fold approximately tetrahedrally coordinated. In the spinel-type γ-structure two thirds of the Si atoms are octahedrally coordinated, the remaining Si are tetrahedrally coordinated.

2 Methods

All first-principles calculation were carried out with the Vienna *Ab initio* Simulation Package (VASP)[3–5]. We used the projector augmented wave (PAW) method[6] in the framework of density functional theory within the generalized gradient approximation according to

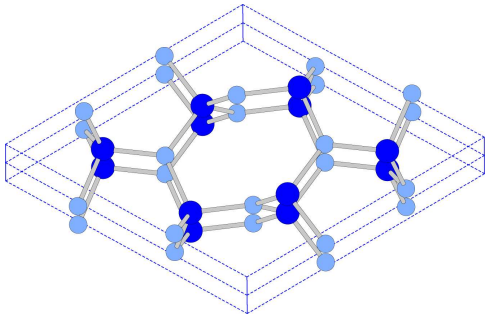


Figure 1. Crystal structure of β-Si_3N_4 ($P6_3/m$). The blue and sky-blue circles denote silicon and nitrogen atoms.

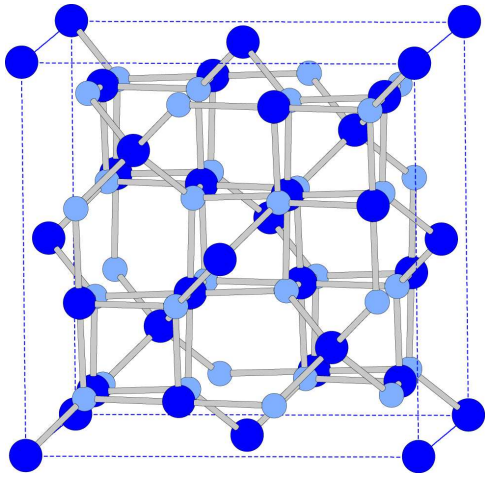


Figure 2. Crystal structure of γ-Si_3N_4 ($Fd\bar{3}m$). The blue and sky-blue circles denote silicon and nitrogen atoms.

Perdew-Burke-Ernzerhof (PBE)[7] together with a cutoff energy of 500 eV for the expansion of the wave function into a plane wave basis set. The phonon contribution to the free enthalpy was calculated using the frozen phonon method as implemented in the *fropho* code[8]. We employed the quasi harmonic approximation (QHA) (e.g.[9]). The free enthalpy like function G(p,T) is computed according to

$$G[p,\,T] = \min_{V} U(V) + F_{\mathrm{phonon}}(V;\;T) + pV, \tag{1}$$

When the investigated phase is a stable phase, then G(p,T) corresponds to its free enthalpy.

Here p, T, and V denote pressure, temperature, and volume, respectively. U is the internal energy, and F_{phonon} is the phonon free energy. We computed the bracketed term at selected volume points. Finally, Murnaghan's equation of state[10] was used to fit the calculated values and to derive a bulk modulus.

To allow for an accurate assessment of the free energy, some heavy convergence crite-

ria have to be employed to avoid the appearance of negative frequencies within the Brillouin zone. Energy was converged until the energy differences in consecutive steps were lower than 10^{-9} eV, and force and stress were converged until they became less than 10^{-6} eV, and 10^{-4} GPa, respectively. The Brillouin zones were sampled by a 448 (β-Si_3N_4) and by a 4x4x4 of (γ-Si_3N_4, primitive unit cell) k-point mesh generated according to the Monkhorst-Pack scheme. In the frozen phonon calculations, a finite atomic displacement of 0.01 Å was used. The supercells were built with 2x2x2 unit cells for β-Si_3N_4 and 2x2x2 primitive cells for γ-Si_3N_4. The meshes for the Brillouin zones were sampled to 8x8x20 and 12x12x12 for β-Si_3N_4 and γ-Si_3N_4, respectively.

3 Results and Discussion

In Figure 3 we show the calculated pressure-temperature (P-T) phase diagram of silicon nitride including the two phases β-Si_3N_4 and γ-Si_3N_4. The difference in the free enthalpy like functions between the two phases, hence the value of ΔG, is represented by the intensity of the colour. The phase boundary corresponds to the line drawn; here the difference in ΔG=0. Accordingly, while β-Si_3N_4 is stable at low pressure and γ-Si_3N_4 at high pressure, the slope of the phase boundary between the two phases is found to be positive. Consequently, higher temperatures will require some higher pressures for inducing the β to γ phase transition.

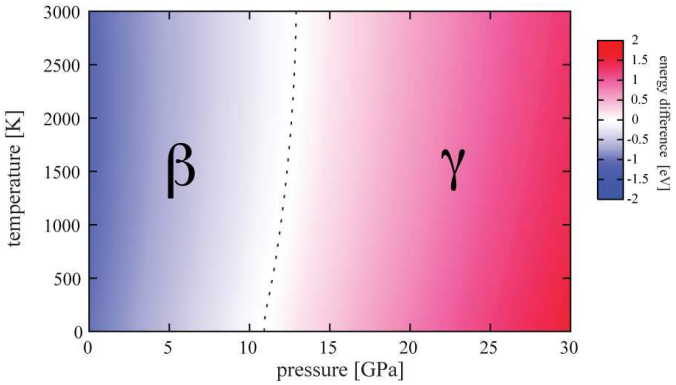


Figure 3. P-T phase diagram. The energies are given per formula unit.

We show the phonon density of states (phonon-DOS) as a function of phonon frequencies of both phases in Figure 4. The phonon-DOS of β-Si_3N_4 is split into two regions. The occurrence of a gap is attributed to the anisotropic crystal structure. Although N is twice as light as Si, nitrogen contributes significantly to the phonon-DOS even in the low-frequency regime. This is due to vibrations in the a-b plane localized at approximately the Brillouin zone boundary. For γ-Si_3N_4 on the other side we don't observe a gap, although a separation into N-dominated high-frequency vibrations and Si-dominated low-frequency vibrations is visible. The compact structure, therefore, drives both regions to overlap considerably.

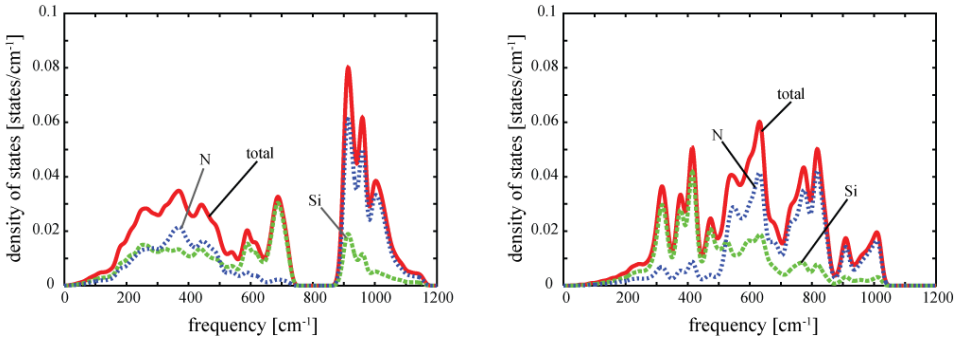


Figure 4. Phonon DOS's of the β-phase (left) and the γ-phase (right).

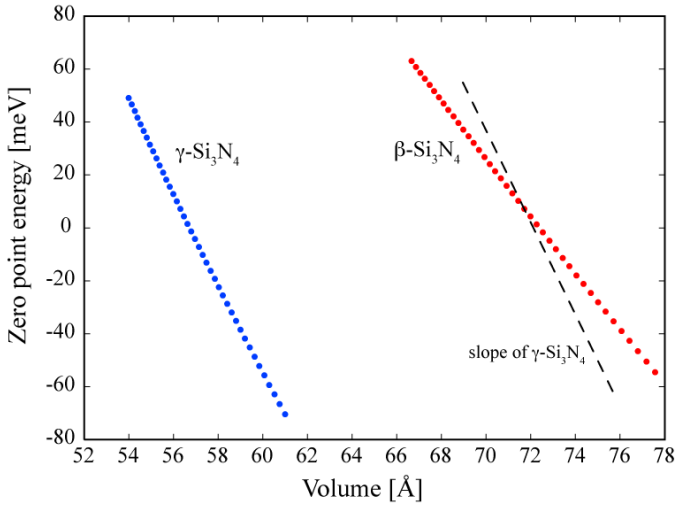


Figure 5. Zero point energies of the β-phase (red points) and the γ-phase (blue points) as a function of volume. To compare their slopes, a copy of γ-phase is placed near that of β-phase. The energies are given per formula unit.

To study the thermal expansion of a solid structure, the volume dependence of its phonon-DOS is a necessary ingredient, since volume and temperature are implicitly related within the QHA. In principle, the phonon free energy also depends directly on temperature. However, in order to investigate the volume dependence of the phonon-DOS only it is usually enough to derive a trend of the phonon free energy with temperature and volume. Therefore, the zero point energy of a phonon system is indicative of how its phonon-DOS changes with volume, because the zero point energy E_{zero} is related to the phonon-DOS via

$$E_{\mathrm{zero}} = \frac{1}{2} \sum_{\mathbf{k},\nu} \hbar\omega(\mathbf{k}, \nu) \tag{2}$$

Here, **k** is a wave vector, and ν a phonon frequency. In Figure 5 we plot the frequency depending part of the zero point energies of β-Si_3N_4 and γ-Si_3N_4 as a function of volume. A negative slope indicates that a volume increase decreases the phonon free energy at 0 K, which is generally expected to be true at finite temperature. However, the data also shows a steeper slope for the curve of γ-Si_3N_4 than for β-Si_3N_4. Consequently, the γ-phase will tend to exhibit a larger thermal expansion than the β-phase.

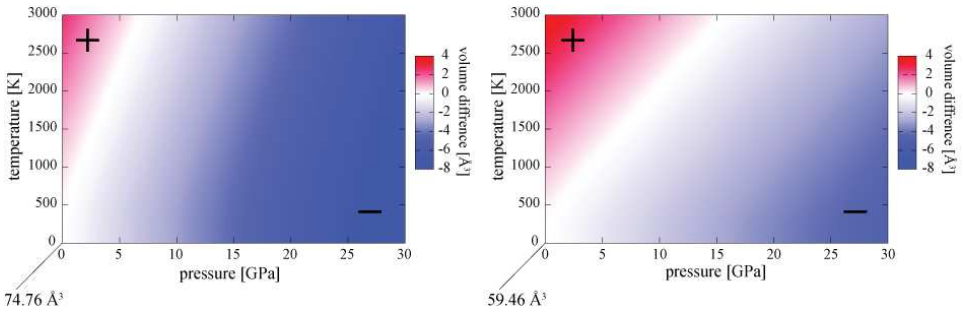


Figure 6. Volume changes of β-Si_3N_4 (left) and γ-Si_3N_4 (right) in the pressure-temperature phase diagram. The change (per formula unit) is indicated relative to their volumes at 0 K and 0 GPa.

The computed volumes of the enthalpy confirm this trend, derived through inspection of the zero point energies. Note, that the enthalpy in turn was assessed as the minimum of computations for several volumes. In Figure 6 we show the overall volume changes of both the β- and the γ-phase in the pressure-temperature space. The thermal expansion of γ-Si_3N_4 is indeed larger than that of β-Si_3N_4. However, at the same time, γ-Si_3N_4 is also less compressible than β-Si_3N_4.

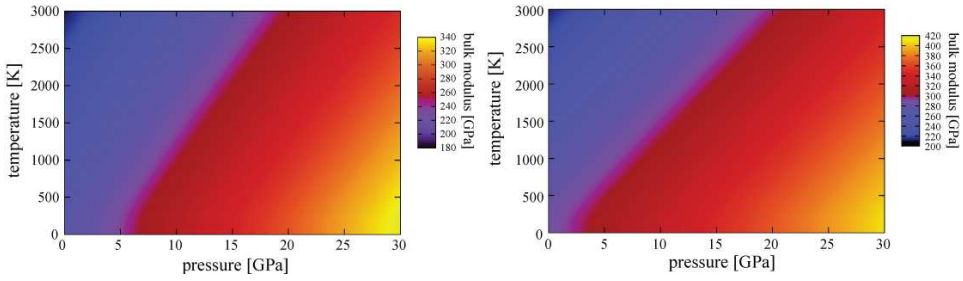


Figure 7. Bulk modulus of β-Si_3N_4 (left) and γ-Si_3N_4 (right) as a function of temperature and pressure.

We present the temperature and pressure dependent bulk modulus, the inverse of compressibility, for β-Si_3N_4 and γ-Si_3N_4 in Figure 7. As is well known, γ-Si_3N_4 has a larger

bulk modulus than β-Si_3N_4. As it turns out, however, temperature affects the bulk modulus of γ-Si_3N_4 more than that of β-Si_3N_4.

4 Conclusions

Using first-principles lattice dynamics calculations to investigate the vibronic contributions to the free enthalpies of phenacite- and spinel-type silicon nitride, we find a positive slope of the phase boundary between β-Si_3N_4 and γ-Si_3N_4 in the pressure-temperature phase diagram. While the spinel-phase has a larger bulk modulus than the phenacite-phase, its thermal expansion is also larger. The reason for this behaviour is the larger volume dependence of the zero point energy in γ-Si_3N_4 in comparison to β-Si_3N_4. Overall, the work of pressure, p·V, dominates the transition pressure of the β to γ transition. The precise slope of the phase boundary, however, is governed by small energy differences, which depend on the explicit phonon structure.

Acknowledgments

Financial support came from the Deutsche Forschungsgemeinschaft (Bonn, Germany) (DFG) within the priority program 1236, Synthesis, “in situ” Characterization and Quantum Mechanical Modeling of Earth Materials, Oxides, Carbides and Nitrides at Extremely High Pressures and Temperatures (contracts KR1805/10-1 and KR1805/11-1). The DFG is also deeply thanked for granting a Heisenberg-Fellowship to P.K. Computations were carried out at the FZ Jülich, which we acknowledge for its generous support.

References

1. Zerr, G. Miehe, G. Serghiou, M. Schwarz, E. Kroke, R. Riedel, H. Fuess, P. Kroll, R. Boehler, Nature 1999, 400, 340.
2. M. Schwarz, G. Miehe, A. Zerr, E. Kroke, B. T. Poe, H. Fuess, and R. Riedel, Adv. Mater. 2000, 12, 883.
3. G. Kresse, J. Non-Cryst. Solids 193, 222 (1995).
4. G. Kresse and J. Furthmller, Comput. Mater. Sci. 1996, 6, 15.
5. G. Kresse and D. Joubert, Phys. Rev. B 1999, 59, 1758.
6. P. E. Blchl, Phys. Rev. B 1994, 50, 17953.
7. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 1996, 77, 3865.
8. A. Togo, `http://fropho.sourceforge.net/`.
9. A. R. Oganov, M. J. Gillan, and G. D. Price, Phys. Rev. 2005, B 71, 64104.
10. F. D. Murnaghan, Proc. Natl. Acad. Sci. USA 30, 244 (1944).

Glycine at the Pyrite–Water Interface: An *ab initio* Metadynamics Study

Nisanth N. Nair, Eduard Schreiner, and Dominik Marx

Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum, D–44780 Bochum
E-mail: {nisanth.nair, eduard.schreiner, dominik.marx}@theochem.rub.de

This research focuses on a prebiotic synthesis scenario of peptides at extreme conditions similar to those near hydrothermal vents. It aims to assess the importance of such high temperature and pressure conditions including mineral surfaces and to gain vital molecular level understanding about these reactions, which is otherwise not possible to obtain in laboratory experiments. The work presented here is about large–scale *ab initio* Car–Parrinello molecular dynamics simulations on characterizing the reactivity of the ideal and the defective pyrite surface having an interface with water at 500 K and about 20 MPa. Desorption mechanisms and free energies for desorption are obtained from *ab initio* metadynamics simulations. Surface defects are found to have an important role in stabilizing the reactant molecules on the surface.

1 Introduction

The quest for a molecular level understanding of chemical reactions and the role of prebiotic conditions in the formation of macromolecules of contemporary biology from small organic molecules is at the heart of our research endeavour. Among the vastly different hypotheses put forward on how life originated on earth, the one suggested by Wächtershäuser[1–3] named 'Iron Sulfur World' (ISW) hypothesis, is under great attention, especially due to exceptionally specific details at the molecular level. Wächtershäuser envisages a chemoautotrophic origin of life and the basic reactions are believed to have taken place near undersea hydrothermal vents. Successful carbon–insertion reactions[4–6] and synthesis of peptides[7,8] in laboratory experiments mimicking hydrothermal conditions brought a strong support to this hypothesis recently[9].

An interesting idea has been suggested for the activation of amino acids that makes use of an inorganic compound called carbonyl sulfide (COS), an effluent gas from hydrothermal vents; alternatively it can be generated from CO and H_2S in the presence of suitable minerals[7]. Though a very small yield of peptides are found in laboratory experiments at elevated temperature, the detailed mechanism or the energetics of these reaction sequences are not known. The mechanistic insights into the formation of NCA from amino acids and COS are limited to the early speculations in footnote (16) of Ref.[10] and some evidence for the presence of an isocyanate species[8].

One reason for the lack of in–depth understanding[11] of these reactions is that the controlled reaction studies in experiments at water/mineral interfaces subject to extreme conditions are extremely challenging. In contrast, conditions which may be hard or even impossible to meet in the laboratory might be accessible 'in silico'. Computational approaches have the benefit that the reactions can be very well controlled and analyzed, in particular the influence of different reaction conditions, for e.g. extreme versus ambient, can be substantiated while keeping all other parameters unchanged. Different chemical reaction pathways can be traced and mechanism as well as free energy barriers involved

can be examined at the molecular level in computer simulations. Recent advances in *ab initio* molecular dynamics[12,13] together with the modern computer infrastructure, like the IBM Blue Gene/L machine 'JUBL' at John von Neumann Institute for Computing at Forschungszentrum Jülich, make it possible to study such complex chemical reactions fully 'in silico'.

In this work we set our aim to look at the reactivity of a wet mineral surface at extreme thermodynamic conditions in our 'virtual lab' approach. In the ISW scenario, it is important that the metabolites stay at the pyrite surface until they concentrate, react and/or undergo structural changes. Thus in the first *ab initio* molecular dynamics simulation at 'ISW conditions' carried out in our laboratory[14] we focused on glycine adsorption at the pyrite–water interface using an ideal (001) surface of FeS_2 (pyrite). It was concluded that water–mediated detachment of glycine occurs on a time scale of picoseconds[14], which is much too fast to allow for chemical reactions to occur there. Starting with a bi–dentate adsorption mode with respect to the carboxylate oxygen atoms of the glycine, Pollet et al.[15] found only partial detachments, but not full desorption within 10 ps time scale. On the other hand, naturally occurring pyrite has plenty of surface defects, predominantly missing sulfur atoms. We have recently shown that such defects increase the reactivity and time scale for desorption[16]. Most importantly, these works underlined the role of hydrogen bonding of the aqueous environment in the desorption processes.

In the presented work, we will look at full desorption of the bidentate adsorption mode considered by Pollet et al.[15]. The detailed desorption mechanism and the associated free energy landscape and barriers are computed employing the *ab initio* metadynamics technique. The results are compared with the desorption at the defective pyrite surface having a single sulfur vacancy[16].

2 System Setup

The mineral surface was modelled by a pyrite (FeS_2) slab consisting of 24 Fe and 48 S atoms. We used an orthorhombic periodic box of dimensions $10.8 \times 10.8 \times 18.9$ Å[14–16]. The supercell contained 36 water molecules corresponding to a density of ≈ 0.85 g/cm^3[17] between two layers of pyrite; see Fig. 1. The pyrite slab exposed the (001) surface towards water mimicking a water/pyrite interface in terms of a standard slab setup. The whole system was thermostated to 500 K resulting in an expected pressure of about 20 MPa and the surface vacancy was created by removing one of the threefold coordinated surface sulfur atoms.

3 *Ab initio* Molecular Dynamics and Metadynamics

Ab initio molecular dynamics[12,13] is well suited for the present project where all calculations were performed within spin–restricted Kohn–Sham density functional theory in its plane wave / pseudopotential formulation[13]. Details are skipped here but are presented in Ref.[15,16,22]. Our previous calculations demonstrated that using this setup the unit cell parameters[15] and the electronic structure[16] are reproduced very well. All calculations presented here were performed with the `CPMD` simulation package[18,13].

Most of the chemical reactions of our concern occur at time scales which are several orders of magnitude beyond those accessible by standard *ab initio* molecular dynamics.

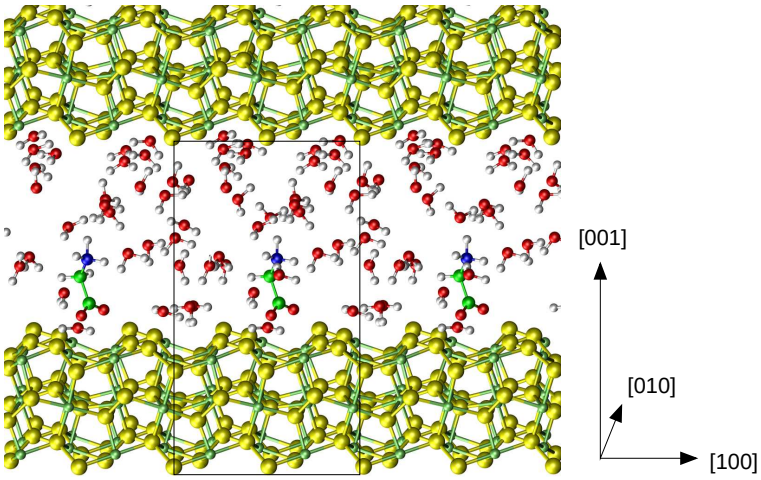


Figure 1. Supercell (black box) used for the simulations of glycine adsorbed at the (001) surface of pyrite in contact with hot–pressurized water; periodic images are included in the [100] and [001] directions to guide the eye. Color code: hydrogen (white), oxygen (red), carbon (green), nitrogen (blue), sulfur (yellow), iron (light green).

For overcoming the timescale limitation by accelerating the sampling of rare events and to estimate free energies and barriers we used the metadynamics technique[19] in its extended Lagrangian formulation[20], see Ref.[21] for a review. At the heart of this technique is a coarse–graining approach where selected degrees of freedom are represented by a small set of collective coordinates[19]. Within this subspace the free energy hypersurface spanned by the collective coordinates is sampled by placing auxiliary Gaussian potentials along the dynamical time evolution. This biasing potential is slowly grown during the 'metadynamics' of the collective coordinates and finally the system is forced to escape from the initial free energy minimum thus exploring reaction pathways. As the added biasing potential compensates the underlying free energy surface the method allows one to map the free energy surface by reconstructing the sum of the added Gaussians. One of the great advantages of this approach is that it enables one to explore not–foreseen transformation and reaction pathways[21].

4 Results

4.1 Desorption of Glycine from the Ideal Surface

In the following we present the desorption study of the bidentate adsorption mode of zwitterionic glycine as sketched in structure **1** of Fig. 2. In earlier investigations of Pollet et al.[15] using short–timescale *ab initio* molecular dynamics it was found that the carboxylate oxygen atoms were partially detaching several times due to hydrogen bonding interactions with water. This behaviour is in stark contrast to that observed in the previous study[14] where a monodentate zwitterionic glycine on the ideal pyrite surface desorbs readily within 1.5 ps. In order to assess the stability of the bidentate adsorption mode in terms of free energy

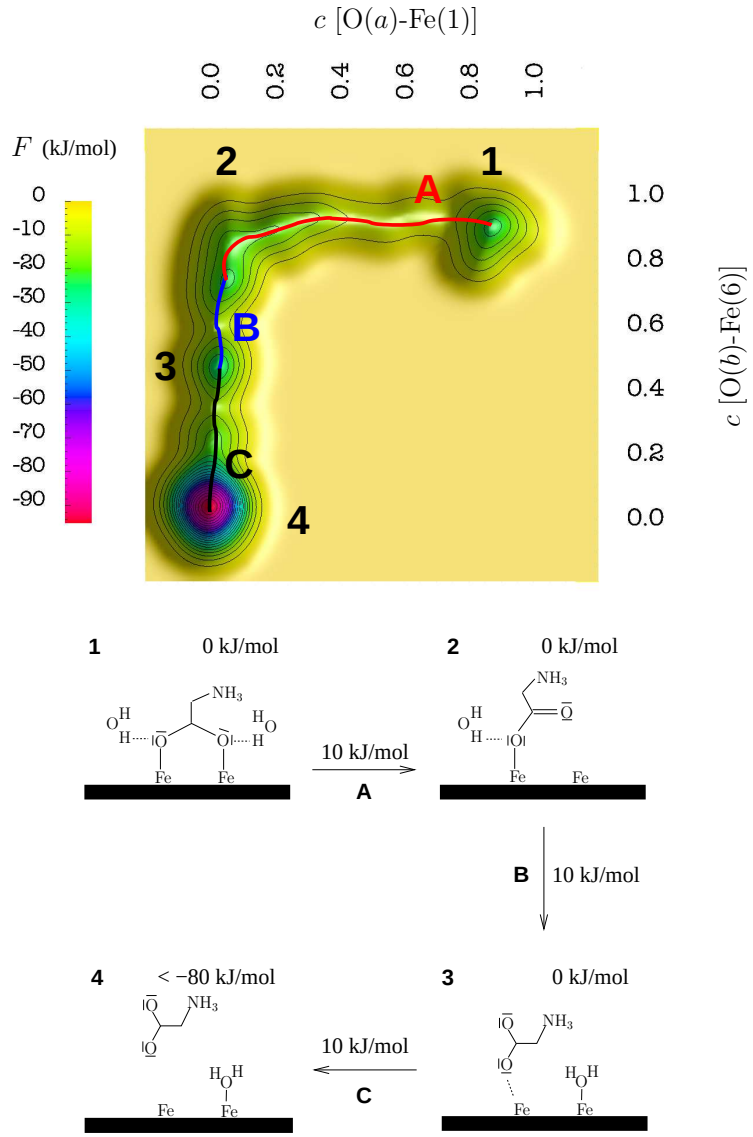


Figure 2. Free energy surface (top) represented in collective coordinate space and corresponding schematic mechanism (bottom) for the desorption process of glycine in the bidentate zwitterionic adsorption mode **1** at the FeS_2(001)–water interface under ISW conditions. Bold numbers and letters denote free energy minima and minimum free energy paths connecting the wells, respectively. The energies in units of kJ/mol associated with structures are the free energies relative to the initial structure **1** and those associated with arrows are free energy barriers.

landscapes and retention times and to investigate the underlying desorption mechanism in pertinent detail we employed *ab initio* metadynamics to enforce the desorption process.

The starting structure **1** of our simulation was the final structure of the simulation by

Pollet et al.[15], i. e. a structure after 10 ps of standard *ab initio* simulations. In the starting structure **1** two carboxylate oxygen atoms are connected to the surface iron atoms and hydrogen bonds exist between water molecules and the carboxylate oxygens. The free energy surface obtained from the metadynamics simulation together with the schematic mechanism along the minimum *free* energy path are given in Fig. 2. Detachment of first carboxylate oxygen occurs forming a monodentate structure **2** by overcoming a free energy barrier of about 10 kJ/mol. Subsequently the O–Fe bond of **2** breaks to form structure **3** which is a partially dissociated structure, held near the surface by the hydrogen bond network of water. The barrier is nearly 10 kJ/mol for the conversion of **2** $\rightarrow$ **3**. For a complete desorption from structure **3** the barrier to surmount is only about 10 kJ/mol. The structures **2** and **3** have almost the same stability as the starting structure within the error bar estimate. The desorbed structure is much lower in free energy compare to all other structures. These results show that the effective barrier for the desorption of bidentate zwitterionic glycine from the ideal surface is only about 10 kJ/mol. In terms of retention time of glycine this means only a few picoseconds, which is again not sufficient for chemical reactions to occur.

4.2 Desorption of Glycine from Defective Surface

In order to design a more realistic mineral surface we have removed one three–fold coordinated surface sulfur atom, which mimics the presence of ubiquitous sulfur vacancies on the pyrite surface. Such defects were found to increase the reactivity compared to an ideal surface, which was probed by both electronic structure calculations and a water adsorption study around the defect site[16]. We have performed a systematic study of various adsorption modes of glycine and extracted two most likely configurations at the pyrite–water interface. The results of a cyclic adsorption mode **5** are presented in Fig. 3.

In particular, different minima and pathways were explored; see Fig. 3. Although the free energy is much lower in the fully solvated (desorbed) state of the glycine molecule, the free energy barriers allowing for desorption are quite high. Compared to the zwitterionic adsorption structure on the ideal surface, the activation barrier for the desorption of the cyclic configuration is close to an order of magnitude higher, 90 kJ/mol. This difference in the activation barriers can be traced back to the fact that the desorption of the zwitterionic form is mediated by hydrogen bonding of solvation water with the carboxylate oxygen, which is not possible for the cyclic adsorption structure due to geometric shielding effects. These results indicate that the cyclic adsorption structure, which can only be formed in the presence of surface vacancies, is stable toward desorption on a millisecond scale at the defective pyrite–water interface. The retention time of the zwitterionic configuration is extended from picoseconds on the ideal pyrite–water interface to nanoseconds in the presence of these point defects.

4.3 Peptide Synthesis

In parallel, we have worked out possible reaction pathways that lead from individual amino acids to peptides and established the role of ISW conditions on these chemical reactions[22]. At the heart of this work is the study of the formation a peptide bond between two glycine molecules using *ab initio* metadynamics techniques. This involved about 27 reaction steps and comprises simulations at three vastly different reaction conditions: ambient bulk water

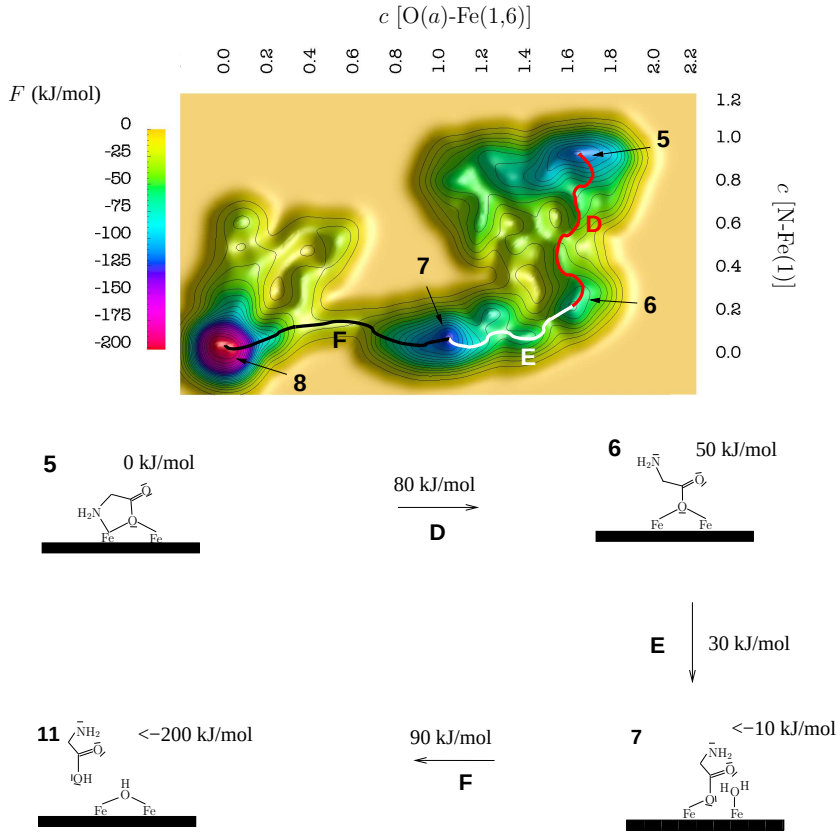


Figure 3. Free energy surface (top) represented in collective coordinate space and corresponding schematic mechanism (bottom) for the desorption process of glycine in the cyclic adsorption mode **5** at the defective $FeS_2(001)$–water interface under ISW conditions. Bold numbers and letters denote free energy minima and minimum free energy paths connecting the wells, respectively. The energies in units of kJ/mol associated with structures are free energies relative to the initial structure **5** and those associated with arrows are free energy barriers.

at about 300 K and 0.1 MPa, hot–pressurized bulk water at about 500 K and 20 MPa, and hot–pressurized water at the pyrite–water interface. These results are already presented in detail in Ref.[22] and are not discussed here.

5 Conclusions

Using extensive *ab initio* metadynamics simulations we have obtained mechanisms and free energy for the desorption of glycine from a hot–pressurized water interface with an ideal and a defective pyrite surface. We have shown that surface defects are necessary for retaining glycine on the surface for long enough times to allow for chemical reactions to occur, which is an essential prerequisite for heterogenous catalysis of prebiotic reactions at mineral surfaces. Access to a supercomputing infrastructure like the IBM Blue Gene/L machine at NIC enabled us to investigate vital primordial chemical reactions and to estimate

their associated free energies employing large scale *ab initio* simulations.

Acknowledgments

Partial financial support was provided by DFG (Normalverfahren MA 1547/7) and by FCI. The simulations were carried out on the IBM Blue Gene/L system 'JUBL' at John von Neumann Institute for Computing at Forschungszentrum Jülich (Germany).

References

1. G. Wächtershäuser, *Before enzymes and templates: theory of surface metabolism*, Microbiol. Rev., **52**, 452–484, 1988.
2. G. Wächtershäuser, *Groundworks for an evolutionary biochemistry: the iron-sulphur world*, Prog. Biophys. molec. Biol., **58**, 85–201, 1992.
3. G. Wächtershäuser, *On the Chemistry and Evolution of the Pioneer Organism*, Chemistry & Biodiversity, **4**, 584–602, 2007.
4. C. Huber and G. Wächtershäuser, *Activated Acetic Acid by Carbon Fixation on (Fe,Ni)S under Primordial Conditions*, Science, **276**, 245–247, 1997.
5. C. Huber, W. Eisenreich, S. Hecht, and G. Wächtershäuser, *α–Hydroxy and α–Amino Acids Under Possible Hadean, Volcanic Origin–of–Life Conditions*, Science, **314**, 630–632, 2006.
6. G. D. Cody, N. Z. Boctor, T. R. Filley, R. M. Hazen, J. H. Scott, A. Sharma, and H. S. Yoder Jr., *Primordial Carbonylated Iron–Sulfur Compounds and the Synthesis of Pyruvate*, Science, **289**, 1337–1340, 2000.
7. C. Huber and G. Wächtershäuser, *Peptides by Activation of Amino Acids with CO on (Ni,Fe)S Surfaces: Implications for the Origin of Life*, Science, **281**, 670–672, 1998.
8. L. Leman, L. Orgel, and M. R. Ghadiri, *Carbonyl Sulfide–Mediated Prebiotic Formation of Peptides*, Science, **306**, 283–286, 2004.
9. G. Wächtershäuser, *From volcanic origins of chemoautotrophic life to Bacteria, Archaea and Eukarya*, Phil. Trans. R. Soc. Lond. B, **361**, 1787–1808, 2006.
10. R. S. Deweye, F. Schoenewald, H. Joshua, W. Paleveda, H. Schwam, H. Barkemeyer, B. H. Arison, D. F. Veber, R. G. Strachan, J. Milkowski, R. G. Denkewaltear, and R. Hirschmann, *The Synthesis of Peptides in Aqueous Medium. VII: The Preparation and Use of 2,5-Thiazolidinediones in Peptide Synthesis*, J. Org. Chem., **36**, 49–59, 1971.
11. C. de Duve and S. L. Miller, *Two–dimensional life?*, Proc. Natl. Acad. Sci., **88**, 10014–10017, 1991.
12. R. Car and M. Parrinello, *Unified Approach for Molecular Dynamics and Density-Functional Theory*, Phys. Rev. Lett., **55**, 2471–2474, 1985.
13. D. Marx and J. Hutter, *Ab initio Molecular Dynamics: Theory and Implementation*, in: Modern Methods and Algorithms of Quantum Chemistry, J. Grotendorst, (Ed.), vol. 3, pp. 301–449, NIC. FZ Jülich, 2000.
14. C. Boehme and D. Marx, *Glycine on a Wet Pyrite Surface at Extreme Conditions*, J. Am. Chem. Soc., **125**, 13362–13363, 2003.

15. R. Pollet, C. Boehme, and D. Marx, *Ab Initio Simulations of Desorption and Reactivity of Glycine at a Water–Pyrite Interface at "Iron–Sulfur World" Prebiotic Conditions*, Origins Life Evol. Biospheres, **36**, 363–379, 2006.
16. N. N. Nair, E. Schreiner, and D. Marx, *Glycine at the Pyrite-Water Interface: The Role of Surface Defects*, J. Am. Chem. Soc., **128**, 13815–13826, 2006.
17. D. A. Lown, H. R.Thirsk, and Lord Wynne–Jones, *Temperature and Pressure Dependence of the Volume of Ionization of Acetic Acid in Water from 25 to 225°C and 1 to 3000 bars*, Trans. Faraday Soc., **66**, 51–73, 1970, Note: the typographical error in the specific volume of 1.1178 cm^3/g for 200 bar pressure and 225°C temperature in Table 5 (page 62) is corrected to 1.1778 cm^3/g.
18. J. Hutter et al., *CPMD*, IBM Corp 1990-2007, MPI für Festkörperforschung Stuttgart 1997-2001, see also `http://www.cpmd.org`.
19. A. Laio and M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci., **99**, 12562–12566, 2002.
20. M. Iannuzzi, A. Laio, and M. Parrinello, *Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics*, Phys. Rev. Lett., **90**, 238302–1–4, 2003.
21. A. Laio and M. Parrinello, *Computing free energies and accelerating rare events with metadynamics*, in: Computer simulations in condensed matter: From materials to chemical biology, M. Ferrario, G. Cicotti, and K. Binder, (Eds.), vol. 1, pp. 315–347, Springer Verlag, Berlin Heidelberg, 2006.
22. E. Schreiner, *Biochemical Aspects of Iron-Sulfur Systems:Magnetostructural Properties of Ferredoxins and Prebiotic Peptide Synthesis Involving Pyrite*, PhD Thesis, Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum, Germany, 2007.

Supramolecular Chemistry from Supercomputers

Werner Reckien[1], **Christian Spickermann**[2], **Thomas Bredow**[1], **and Barbara Kirchner**[2]

[1] Institut für Theoretische und Physikalische Chemie
Universität Bonn, Wegelerstraße 12, D-53115 Bonn
E-mail: werner@thch.uni-bonn.de

[2] Wilhelm-Ostwald Institut für Physikalische und Theoretische Chemie
Universität Leipzig, Linnestraße 2, D-04103 Leipzig

We present the results of a computational study of different supramolecular systems which are assumed to enable the access to molecular machines. Mainly we treat systems which are involved in the formation of rotaxanes. Rotaxanes constitute a class of supramolecular complexes consisting of two mechanically linked components, a macrocyclic wheel and a linear molecule (axle). A study of the dynamical properties of a tetralactam macrocycle shows its ability to adapt to potential guest molecules. A constrained simulation of the dethreading of a pseudorotaxane demonstrates that the breaking and formation of hydrogen bonds is essential for this process. An analysis of different binding situations for the pseudorotaxane indicates that dispersion interaction may also play a role in this context. In addition we can show that it should be possible to distinguish different binding situations of pseudorotaxane complexes by means of IR spectroscopy. We show that hydrogen bonds are also responsible for the observed self organization of TLM on the Au(111) surface. Finally we present first results of the investigation of peptide hydrogen bonds on the Ag(111) surface.

1 Introduction

Supramolecular chemistry comprises chemical processes of large molecules often interacting via many, comparatively weak, non bonding contacts. An important and generic process in supramolecular chemistry is molecular recognition in host-guest complexes, which is for instance, the basis of protein-substrate interactions in biochemistry. A few examples for supramolecular systems are rotaxanes, catenanes, molecular knots, calixarenes, and molecular tweezers.[1]

Here we present a computational study of supramolecular systems which contain a tetralactam macrocycle (TLM) as central motif, see Fig. 1. The TLM can act as receptor for p-benzoquinone[2] and is widely used for the synthesis of catenanes and rotaxanes.[3] In addition it can act as a sensor for the detection of trans-2-hexenal and acrylamide.[4] The special features of the TLM are its ability to build up a multiplicity of different hydrogen bonds (The TLM has four amide groups which can act either as hydrogen bond donor or as acceptor.) and its comparative rigid cavity. It is assumed that the interaction between the TLM host and the guest can be controlled by external influences. This enables the access to molecular machines.

One example for these systems are rotaxanes. Rotaxanes constitute a class of supramolecular complexes consisting of two mechanically linked components which are referred to as axle, and wheel. Sterically demanding functional groups at the ends of the axle prevent a dethreading. In this work we investigated so-called pseudorotaxanes where the wheel is threaded on an axle omitting appropriate stopper groups. The resulting complex is linked via non-covalent interactions like hydrogen bonds.

Furthermore we studied in cooperation with the experimental group of M. Sokolowski the adsorption of hydrogen bond templated supramolecular systems. The adsorption of building blocks of molecular machines is interesting since it constitutes a first step towards the design of 2D functional arrays of these molecules. We show that a hydrogen bond mediated template effect can lead to the formation of ordered 2D structures of larger and non planar molecules on surfaces. Hereby a specific functionalization of surfaces is possible.

This article is organized as follows: We start with a short description of the methodology we use for the investigation of large supramolecular systems which is followed by a short summary of the investigation of the dynamical properties of the TLM and of the dethreading process of a pseudorotaxane complex. Next we show the results of an energetic and vibrational analysis of different binding situations of the pseudorotaxane system. Finally we present the hydrogen bond mediated self assembly of the TLM on the Au(111) surface and the adsorption of formamide polymers on the Ag(111) surface.

2 Methodology

Density functional theory (DFT)[5] is nowadays the most-widely used electronic structure method. Since DFT combines reasonable accuracy in different chemical environments with minimal computational effort we chose this method for the investigation of large supramolecular systems. The calculations presented in this article are in large part performed with plane wave DFT codes, namely the Car-Parrinello Molecular Dynamics code (CPMD)[6] and the VASP code[7] which are well suited for an efficient parallelization. The numerical effort of applying a DFT plane wave code mainly consists of basic linear algebra subprograms (BLAS) and fast Fourier transform (FFT) operations. The former ones generally require quite little communication. However the latter one requires more complicated communication patterns since in larger systems the data on which the FFT is performed needs to be distributed on the processors. Combined with a suitable grouping of the FFT's one can achieve good scaling up to tens of thousands of processors.

3 Results

3.1 Dynamical Properties of Macrocycles

We started our study with a detailed investigation of the isolated tetralactam macrocycle TLM, see Fig. 1. For this purpose we performed a Car-Parrinello molecular dynamics simulation with a time step of 0.12 fs. From the total trajectory of 1.5 ps we harvested every third time step the Wannier functions. During the simulation we observe a conformational changing from the so called 4-out-0-in conformation to the 3-out-1-in conformation (out: C=O (hydrogen bond acceptor) pointing away from cavity, N-H (hydrogen bond donor) pointing inside the cavity; in: C=O pointing inside cavity). This observed conformational flexibility of the otherwise rigid TLM is connected with a changing of the TLM ability to build up hydrogen bonds to guest molecules. This demonstrates that the TLM can easily adapt its properties to potential guests.

Since Diederich and coworkers have pointed out the importance of inter- and intramolecular multipole interactions in supramolecular chemistry[8] we turn our attention to changes of the total dipole moment during the simulation. In Fig. 1 we show the dipole

along the time step. We observe in the right graph that the dipole drops until the 750th time step is reached. After this the dipole moment suddenly rises. This change of dipole moments could influence the intra-supermolecule dynamics of wheel and axle in a rotaxane and thus the potential motion of a rotaxane.

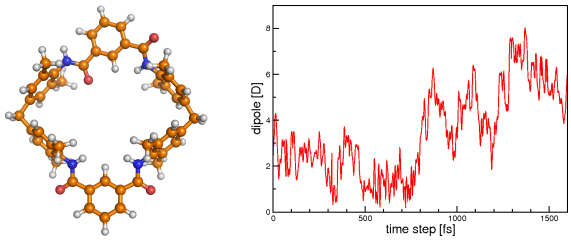


Figure 1. Left: The tetralactam macrocycle (TLM), right: dipole of the TLM during the CPMD simulation.

3.2 Dethreading of a Pseudorotaxane

The aim of our study is to get a deeper insight in the mechanism of the threading and dethreading process of rotaxanes. This knowledge will be useful for the realization of rotaxane based molecular machines. The system under examination is a pseudorotaxane system where a secondary amide axle builds up three hydrogen bonds to TLM, see Fig. 2. We performed constrained CPMD simulations with a time step of 0.12 fs where we used the GROWTH algorithm to enforce the dethreading of the axle. Hereby we enlarged the distance between selected axle and wheel atoms by 0.001 a.u. per time step.

We start the simulation at the most stable conformation of the pseudorotaxane where the axle builds up all together three hydrogen bonds to the wheel. We observe during the simulation the breaking and formation of hydrogen bonds. This process ends in a situation where the axle is connected to the wheel via a single hydrogen bond. This hydrogen bond is retained during the further dethreading process: The corresponding TLM amide group performs a rotation from an *in* to an *out* conformation. This leads to an arrangement where the axle is situated above the macrocycle and not inside the cavity. Although this constrained dethreading is connected with an unphysical deformation of the components one can see that the breaking and formation of hydrogen bonds is essential for this process. In particular one can assume that the first step of the threading process is the formation of a single hydrogen bond outside the cavity.

3.3 Energetic and Vibrational Analysis of Pseudorotaxanes

The results of the constrained dethreading process encouraged us to investigate the different binding situations for the pseudorotaxane in more detail. We performed an energetic and vibrational analysis of complexes which differ in the wheel conformation and in the hydrogen bonding pattern provided by the mutual interactions between the wheel and the axle.[9] We studied a complex where the maximum of three hydrogen bonds is formed (**1a**, see Fig. 2), a complex where the wheel and the axle are connected via a twofold hydrogen

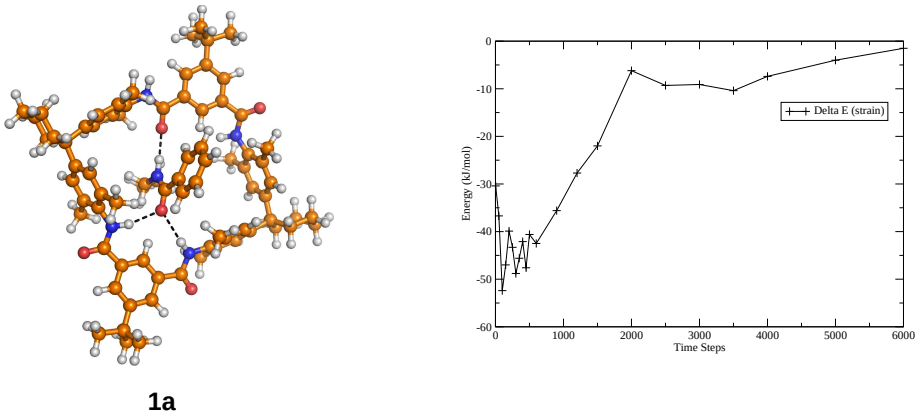


Figure 2. Left: pseudorotaxane complex **1a** with three hydrogen bonds. Right: Interaction energy in kJ/mol during the constrained dethreading (single point calculation of snapshots).

bond (**1b**) and two complexes with only a single hydrogen bond, see Fig 3. For the latter we distinguish between the situation where the axle is placed inside the cavity (**1c**), and the situation where the axle is outside the cavity (**1d**). These chosen minimum structures may serve as different steps of the dethreading event of the pseudorotaxane.

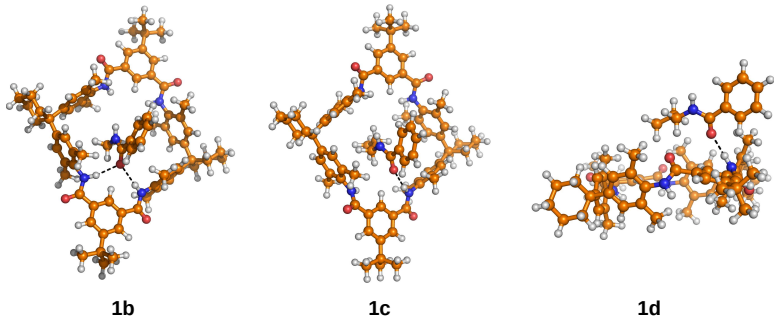


Figure 3. Left: system **1b** (twofold hydrogen bridge), middle: system **1c** (single hydrogen bridge, inside the cavity), right: system **1d** (single hydrogen bridge, outside the cavity)

The interaction energies of these systems are calculated with different density functional methods (BP, BLYP, B3LYP, BHLYP, BLYP-D, and B2-PLYP) as well as with the SCS-MP2 and MP2 method, see table 1. The B2-PLYP and the SCS-MP2 calculations are performed at truncated geometries where the TLM t-butyl and cyclohexyl groups are replaced by hydrogen atoms. This approach enables for the first time a quantitative estimation of the role of dispersion interaction for the formation of such systems.

The most stable structure in these isolated system considerations is one with three hydrogen bonds between axle and wheel. For almost all methods the next stable compound is represented by the structure with the twofold hydrogen bond. This complex is predicted

	BP86	BLYP	B3LYP	BHLYP	BLYP-D	B2-PLYP	SCS-MP2	MP2
1a	-47	-42	-58	-75	-132	-86	-102	-122
1b	-21	-16	-31	-46	-103	-59	-80	-98
1c	-12	-9	-21	-34	-87	-48	-66	-84
1d	-22	-20	-27	-35	-59	-40	-46	-54

Table 1. Interaction energies (in kJ/mol) for the investigated pseudorotaxanes **1a** – **1d**.

to be about 30 kJ/mol above the most stable structure considering all studied functionals. For almost all employed functionals the structure with the axle inside the wheel and one single hydrogen bond lies approximately another 10 kJ/mol above the second most stable structure. The comparison of the two single hydrogen bonded system **1c** and **1d** shows that **1d** is more stable for the standard density functionals with a poor description of dispersion interaction whereas **1c** is more stable for the methods where dispersion interaction is considered in a better way.

It attracts attention that the calculated interaction energy for the three systems with the axle inside the wheel is functional dependent. The trend for the pure density functionals (BP to BHLYP) corresponds to their expected description of hydrogen bonds: BP and BLYP tend to underestimate, BHLYP to overestimate hydrogen bonds. It is assumed that B3LYP provides good results for hydrogen bonds. A large increase of the interaction energy is observed for the other methods. The interaction energy compared to the B3LYP results increases about -28 kJ/mol for B2-PLYP, about -46 kJ/mol for SCS-MP2, about -65 kJ/mol for MP2, and about -71 kJ/mol for BLYP-D. These results indicate that the influence of dispersion interaction is unexpectly large. Therfore one can assume that the first step of the threading process is the formation of a hydrogen bond outside the cavity. The driving power for the continuing threading process is the formation of further hydrogen bonds inside the cavity and the dispersion interaction which is also largest inside the cavity. However one should keep in mind that the dispersion interaction of solvent molecules which are not treated in this work is expected to be of the same order of magnitude as the dispersion interaction of the axle.

The vibrational analysis of the pseudorotaxane complexes demonstrates that it should be possible to distinguish the treated hydrogen bonding situations by means of IR spectroscopy.[9] In particular the N-H stretching modes of the wheel might act as indicator for the binding situation of the axle, since a large difference in the shifts is predicted if we compare the situation where the axle is bound inside the wheel cavity to the situation where the axle is bound outside the cavity. In the latter case we observe by far the largest red shift. The different hydrogen bonding situations for the cases where the axle is positioned inside the cavity can be distinguished on the basis of the carbonyl C=O stretching modes which show also a red shift if they are involved in hydrogen bonds. This red shift corresponds to the hydrogen bonding situation of the carbonyl group. The red shift is larger for a twofold hydrogen bonded C=O group than for a single hydrogen bonded group.

3.4 Adsorption of TLM on the Au(111) Surface

In cooperation with the experimental group of M. Sokolowski we investigated the template driven self organization of TLM on the Au(111) surface.[10] From scanning tunneling microscopy and our calculations two structural different phases with fully flat lying macrocycles were identified. The inert Au(111) surface acts only as a two dimensional template. The intermolecular ordering of the macrocycles results from hydrogen bonds between amide groups of next neighbour macrocycles. The so called α phase, which is apparently the thermodynamically more stable phase at room temperature (Fig. 4) consists of linear chains of stacked TLM molecules.

From DFT calculations we find that two hydrogen bonds between amide groups are formed between each pair of overlapping molecules. In addition $\pi - \pi$-stacking interaction can stabilize the chain formation, since the distance between two overlapping benzene rings is only around 3.7 Å. The structure of the β phase is assumed as linear aggregates of three TLM molecules, see Fig. 4. Again the formation of linear hydrogen bonds is the central structural motive in this phase. Both, the α phase and the β phase feature the maximum possible number of hydrogen bonds.

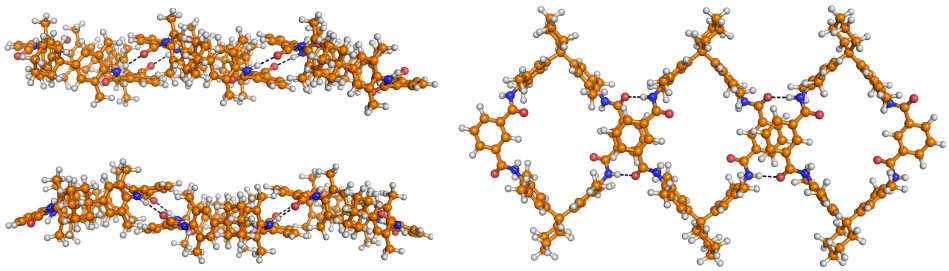


Figure 4. Calculated structure models for the α (left upper and right structure) and β (left lower structure) phase of TLM.

3.5 Adsorption of Formamide Polymers on Ag(111)

In a further project we investigate the influence of surfaces on peptide hydrogen bond templated systems. For this purpose we studied the adsorption of formamide monomers and of formamide polymers in which each formamide molecule builds up to four hydrogen bonds to its neighbours. If these chains are placed on the Ag(111) surface we expect template induced chirality.

We observe a pronounced difference between the adsorption of formamide monomers which corresponds to a low coverage situation and the adsorption of formamide polymers: The adsorption of a monomer leads to a structure where the carbonyl group points downwards to the surface. The calculated adsorption energy is -25 kJ/mol. In contrast the formamide molecules of the adsorbed polymers, which resemble the most stable configuration in the gas phase, are lying flat on the surface. For this system we get significantly larger $\mathrm{R_{OAg}}$ distances and we calculate smaller adsorption energies (-3 to -8 kJ/mol per molecule). However, due to the maximum number of intermolecular hydrogen bonds

this arrangement is energetically favored compared to independently adsorbed monomers. The total stabilization energy due to adsorption and hydrogen bonds is -85 kJ/mol for the dimer, -216 kJ/mol for the tetramer, and -333 kJ/mol for the hexamer. The total interaction energy per molecule increases from -43 kJ/mol (dimer) to -54 kJ/mol (tetramer) to -56 kJ/mol (hexamer) to -58 kJ/mol for the adsorption of an infinite chain of formamide molecules which was also investigated.

A more detailed analysis shows that the formation of hydrogen bonds has an effect on the adsorption energy: The adsorption energy of the polymer is significantly lower than the sum of the adsorption energies of the corresponding monomers. From this follows that the interaction between the formamide and the Ag is weakened by the formation of hydrogen bonds parallel to the surface. It appears that first and foremost the outer oxygens which build up only one single hydrogen bond to their neighbours are responsible for the stabilization due to adsorption.This assumption is confirmed by the fact that we observe shorter $\mathrm{R_{OAg}}$ distances for these atoms.

We studied also the vertical adsorption of a formamide dimer where both carbonyl groups point downwards to the surface. This leads to the formation of a hydrogen bond of one carbonyl group to the adjacent formamide molecule. The calculated stabilization energy (adsorption and hydrogen bonds) for this system is -74 kJ/mol, 10 kJ/mol less than for the flat system. Interestingly we observe for this vertical adsorption the opposite trend to the flat adsorption: The adsorption energy increases due to the formation of a vertical hydrogen bond. Our preliminary conclusion is that the formation of peptide hydrogen bonds parallel to the surface leads to a decrease of adsorption energy wheres the formation of hydrogen bonds vertical to the surface increases the adsorption energy.

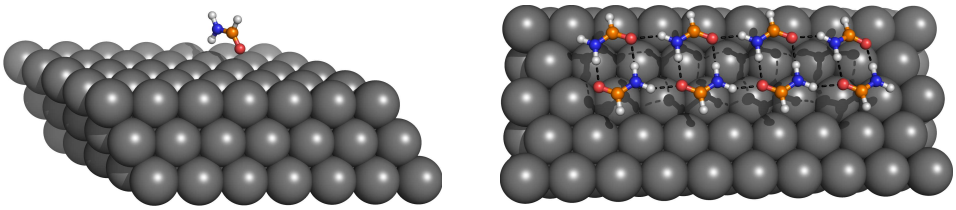


Figure 5. Adsorption on the Ag(111) surface, left: vertical adsorption of formamid monomer, right: flat adsorption of formamid oktamer

4 Computational Performance

Before starting our calculations we measured the scaling of the computing time when changing the accuracy and/or the number of processors incorporated in a job. Our test system for the CPMD simulations was the dethreading process of the pseudorotaxane system **1a**. The results of these calculations are shown in Fig. 6. We observe very good scaling.

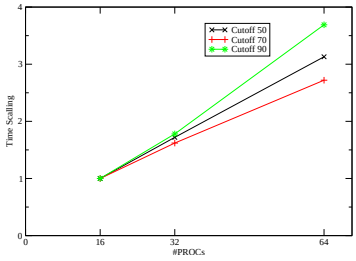


Figure 6. Scaling of the wall clock time for the CPMD simulations of pseudorotaxane **1a**

Acknowledgments

We are grateful for computing time at the NIC in Jülich. Furthermore we acknowledge the financial support from the collaborative research center SFB 624 “Templates” at the University of Bonn, and from the ERA project “Modular Approach to Multi-responsive Surfactant/Peptide (SP) and Surfactant/ Peptide/Nanoparticle (SPN) Hybrid Materials”. We thank our experimental collaborators I. Kossev and M. Sokolowski.

References

1. V. Balzani, M. Venturi, A. Credi *Molecular Devices and Machines* (Wiley-VCH, Weinheim, Germany, 2003).
2. C. A. Hunter, D. H. Purvis *A Binary Quinone Rezeptor* Angew. Chem. Int. Ed. Engl. **31**, 792–795, 1992.
3. C. A. Schalley, W. Reckien. S. D. Peyerimhoff, B. Baytekin, F. Vögtle *Theory and Experiment in Concert: Templated Synthesis of Amide Rotaxanes, Catenanes, and Knots* Chem. Eur. J. **10**, 4777–4789, 2004.
4. U. Herrmann, T. Jonischkeit, J. Bargon, U. Hahn, Q. Y. Li, C. A. Schalley, E. Vogel, F. Vögtle *Monitoring apple flavor by use of quartz microbalances* Anal. Bioanal. Chem. **372**, 611–614, 2002.
5. R. G. Parr, W. Yang *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, USA, 1989).
6. CPMD, see `http://www.cpmd.org`.
7. VASP, see `http://cms.mpi.univie.ac.at/vasp/`.
8. R. Paulini, K. Müller, D. Diederich *Orthogonal multipolar interactions in structural chemistry and biology* Angew. Chem. Int. Ed. Engl. **44**, 1788–1805, 2005.
9. W. Reckien, C. Spickermann, M. Eggers, B. Kirchner *Theoretical energetic and vibrational analysis of amide-templated pseudorotaxanes*, Chem. Phys. in press (doi:10.1016/j.chemphys.2007.09.027).
10. I. Kossev, W. Reckien, B. Kirchner, T. Felder, M. Nieger, C. A. Schalley, F. Vögtle, M. Sokolowski *Highly Ordered 2D Hydrogen-Bonded Structures of a Tetralactam Macrocycle on the Au(111) Surface*, Adv. Funct. Mater. **17**, 513–519, 2007.

Elementary Particle Physics: Lattice Gauge Theories - Broader and Deeper

Richard Kenway

School of Physics, The University of Edinburgh
Edinburgh EH9 3JZ, UK
E-mail: r.d.kenway@ed.ac.uk

The approximation of a finite space-time volume by a regular hypercubic lattice enables quantum field theories in imaginary time to be simulated on a computer as if they were described by equilibrium statistical mechanics. In principle, provided the resulting lattice action is real so that the Boltzmann weight is positive, Monte Carlo methods are very effective at computing expectation values of products of fields. Then, with sufficient computer power, a quantum field theory may be simulated for any choice of its input parameters (masses and couplings for the elementary fields) and the lattice approximation can be controlled systematically by adjusting the lattice spacing while holding the volume fixed and large enough. This approach has given particle physicists a tool for understanding quantum field theories from first principles, without having to rely on perturbation theory, and a laboratory for exploring different parameter values from those in the real world.

The introduction of the lattice breaks some of the symmetries of the continuum quantum field theory. Therefore, it is crucial to demonstrate that the correct symmetries are recovered in the continuum limit, which is reached in practice by extrapolating results from simulations at several lattice spacings. In the language of critical phenomena, this limit is achieved at a point where the lattice theory has a continuous phase transition and physical length scales diverge in units of the lattice spacing. The lattice theory should lie in the same universality class as the continuum theory. Happily, Lorentz symmetry is an accidental symmetry of the lattice theory, i.e., it is only broken by irrelevant operators, and so is automatically recovered in the continuum limit. Ken Wilson's breakthrough in 1974[1] was to discover how local gauge symmetries could be maintained exactly on a lattice (at each of the sites), thereby, opening up the possibility that the confinement of quarks in Quantum Chomodynamics (QCD), which is believed to be due to non-abelian gauge invariance, might be studied and even understood using lattice QCD. Combined with Monte Carlo methods introduced by Michael Creutz in 1980[2], this launched a world-wide programme of QCD simulations.

One symmetry proved problematic – the chiral symmetry of massless quarks, which is spontaneously broken resulting in massless Goldstone pions. In the real world, u and d quarks have small masses which generate correspondingly small pion masses described by chiral perturbation theory. Early lattice formulations of QCD could not avoid explicitly breaking chiral symmetry while preserving the correct flavour content. Now we understand how to realise the full chiral and flavour symmetries at non-zero lattice spacing, but at a higher computational cost, which arises from the fact that chiral symmetry is naturally realised non-perturbatively in one extra space dimension. Having all the internal symmetries

of continuum QCD means that operator mixing and renormalisation are no more complicated than in continuum perturbation theory and this has extended the range of quantities that can be computed reliably. In addition to this theoretical progress, numerical algorithm improvements have delivered at least as great a speed-up as Moore's Law over the past three decades.

Consequently, lattice field theory is currently undergoing both a broadening of its range of application beyond QCD and a deepening of its penetration into QCD. The twisted mass formulation of lattice QCD[3] is a compromise between preserving continuum symmetries and speed that is enabling simulations to go deeper into the regime where they can be matched to chiral perturbation theory. It is then possible to extract low energy constants of this effective theory from which a range of physical quantities may be computed. Algorithmic improvements are enabling simulations of SU(3) Yang-Mills theory with one fermion flavour[4], which has a non-positive definite Boltzmann weight and a very different realisation of chiral symmetry from QCD. These are beginning to test conjectures about the relationship to supersymmetric Yang-Mills theories in the limit of a large number of colours. The overall advance in simulation capability means that lattice QCD is now providing detailed information about the internal structure of the nucleon[5]. Finally, the full power of the chirally-symmetric formulation gives us hope that the 30 year old puzzle of the $\Delta I = 1/2$ Rule can be solved[6]. These results show that lattice gauge theories have become an indispensable tool for a detailed understanding the Standard Model and for insight into the physics that may lie beyond.

References

1. K.G. Wilson, *Confinement of Quarks*, Phys. Rev. D10, 2445-2459, 1974.
2. M. Creutz, *Asymptotic-Freedom Scales*, Phys. Rev. Lett. 45, 313–316, 1980.
3. K. Jansen, *Maximally Twisted Mass Fermions: Towards the Chiral Regime of Lattice QCD*, this volume, pp. 89–96.
4. F. Farchioni, I. Montvay, G. Münster, E. Scholz, T. Sudmann and J. Wuilloud, *Hadron Spectrum of QCD with One Quark Flavour*, this volume, pp. 97–104.
5. G. Schierholz, *The Transverse Spin Structure of the Nucleon*, this volume, pp. 105–111.
6. P. Hernández, M. Laine, B. Knippschild, C. Pena, T. Tilli, E. Torró, J. Wennekers and H. Wittig, *Exploring the $\Delta I = 1/2$ Rule in Non-Leptonic Kaon Decays*, this volume, pp. 113–120.

Maximally Twisted Mass Fermions: Towards the Chiral Regime of Lattice QCD

Karl Jansen

DESY, Zeuthen
Platanenallee 6, D-15738 Zeuthen, Germany
E-mail: Karl.Jansen@desy.de

We describe results from large scale lattice simulations employing maximally twisted mass fermions. With this approach to lattice QCD small values of the quark mass can be reached which allow for a safe contact with chiral perturbation theory. We show that very precise values for the low energy constants of the effective chiral Lagrangian can be obtained in this way.

1 Introduction

The theory to describe the strong interactions, which is responsible e.g. for the existence of the proton and the neutron, is Quantum Chromodynamics. It postulates the existence of fundamental particles, quarks and gluons, which bind together to form the observed hadron spectrum.

The main problems with QCD are first that the fundamental building blocks, the quarks and gluons, cannot be observed themselves, a phenomenon that is called confinement. Secondly, the binding energy of the hadrons is so large that any perturbative treatment of describing them must fail.

A conceptually very attracting way to study QCD is by means of numerical simulations. To this end, the theory is formulated on a discrete space-time lattice with a certain lattice spacing a. Then the system can essentially be regarded as a statistical physics model and the techniques from statistical mechanics can be employed. Indeed, considering simpler models than QCD, this approach to address quantum field theories has been very successful. However, the application to lattice QCD has turned out to be extremely demanding and in the past a number of approximations had to be made to tackle the problem.

This situation has changed in the last two years dramatically. New algorithms have been developed that allow an order of magnitude faster simulations and modern supercomputer architectures are available now that can reach tens and hundreds of teraflops, such as the one installed at NIC. In this drastically improved situation, it is now possible to really perform realistic lattice QCD simulations, leaving all approximations behind and treating the complete theory.

An important aspect is that the results of such simulations which are obtained at non-vanishing values of the lattice spacing have to be extrapolated to the continuum limit at zero lattice spacing. The original formulation of lattice QCD by K. Wilson[1] has discretization effects that are linear in the lattice spacing. This fact has turned out in practical simulations to be rather problematic since it leads to a slow convergence to the continuum limit.

However, it has been possible to develop alternative formulations of lattice QCD where the effects that are linear in the lattice spacing are completely canceled – we speak of $\mathcal{O}(a)$-improvement – and consequently, the continuum limit can be reached much faster. In this

report, we will address one of these approaches, the so-called twisted mass formulation[2,3]. It has to be stressed that there are alternative approaches such as clover improved Wilson fermions and chiral invariant overlap and chirally improved hypercube or truncated perfect action fermions, see e.g. ref.[4] for a discussion and references to these different kind of lattice actions. It is important that different lattice fermions are used in lattice QCD since they have different kind of systematic errors. Using different lattice discretizations of QCD and performing a corresponding continuum limit provides a most valuable test of physical results that come from lattice QCD.

The results that are shown here were obtained within a large European Twisted Mass Collaboration (ETMC) which comprises a number of different places in Europe, i.e. Cyprus: Univ. of Nikosia; France: Univ. of Paris Sud and LPSC Grenoble; Germany: Humboldt Univ. zu Berlin, Univ. of Münster, DESY in Hamburg and Zeuthen; Great Britain: Univ. of Glasgow and Liverpool, Italy Univ. of Rome I, II and III, ECT* Trento; Netherlands: Univ. of Groningen; Spain: Univ. of Valencia; Switzerland: Univ. Zürich. This collaboration brings together a large amount of different expertise, ranging from code optimization and algorithm development to physical applications that are directly relevant for ongoing experiments and phenomenological analyses. A first account of our results can be found in ref.[5]. Overviews can be found in[6,7].

2 Choice of Lattice Action

What we will consider in this contribution is the Wilson twisted mass fermionic lattice action for two flavours of degenerate quarks which reads (in the so called twisted basis[2] and fermion fields with continuum dimensions)

$$\begin{aligned} S_{\rm tm} = a^4 \sum_x \Big\{ \bar{\chi}_x \left[m_0 + i\gamma_5\tau_3\mu + \frac{4r}{a} \right] \chi_x \\ + \frac{1}{2a} \sum_{\nu=1}^{4} \bar{\chi}_x \left[U_{x,\nu}(-r+\gamma_\nu)\chi_{x+\hat{\nu}} + U^\dagger_{x-\hat{\nu},\nu}(-r-\gamma_\nu)\chi_{x-\hat{\nu}} \right] \Big\} . \end{aligned} \tag{1}$$

Here am_0 is the bare untwisted quark mass and $a\mu$ the bare twisted mass, τ_3 is the third Pauli matrix acting in flavour space and r is the Wilson parameter, which we set to one in our simulations. Twisted mass fermions are said to be at *maximal twist* if the bare untwisted mass is tuned to its critical value, $m_{\rm crit}$. We will discuss later how this can be achieved in practice.

In the gauge sector we use the so called tree-level Symanzik improved gauge action (tlSym)[8], which includes besides the plaquette term $U^{1\times1}_{x,\mu,\nu}$ also rectangular (1×2) Wilson loops $U^{1\times2}_{x,\mu,\nu}$

$$S_g = \frac{\beta}{3} \sum_x \left(b_0 \sum_{\substack{\mu,\nu=1 \\ 1\le\mu<\nu}}^{4} \{1 - \mathrm{Re}\,\mathrm{Tr}(U^{1\times1}_{x,\mu,\nu})\} + b_1 \sum_{\substack{\mu,\nu=1 \\ \mu\neq\nu}}^{4} \{1 - \mathrm{Re}\,\mathrm{Tr}(U^{1\times2}_{x,\mu,\nu})\} \right) \tag{2}$$

with β the bare inverse coupling, $b_1 = -1/12$ and the (proper) normalisation condition $b_0 = 1 - 8b_1$. Note that at $b_1 = 0$ this action becomes the usual Wilson plaquette gauge action.

2.1 $\mathcal{O}(a)$ Improvement

As mentioned before, $\mathcal{O}(a)$ improvement can be obtained by tuning Wilson twisted mass fermions to maximal twist. In fact, it was first proved in Ref.[3] that parity even (and therefore physical) correlators are free from $\mathcal{O}(a)$ lattice artifacts at maximal twist by using spurionic symmetries of the lattice action. Later on it was realised[9,10] that a simpler proof is possible based on the parity symmetry of the continuum QCD action and the use of the Symanzik effective theory.

From this latter way of proving $\mathcal{O}(a)$ improvement, it becomes also clear how to define maximal twist: first, choose an operator odd under parity (in the physical basis) which has a zero expectation value in the continuum. Second, at a non-vanishing value of the lattice spacing tune the expectation value of this operator to zero by adjusting the value of am_0. This procedure, which has been proposed in[11,12] and has been theoretically analysed in[9], is sufficient to define maximal twist independently of the chosen operator. To approach smoothly the continuum limit this tuning has to be performed at fixed physical situation while decreasing the lattice spacing.

Besides being a theoretically sound formulation of lattice QCD, Wilson twisted mass fermions offer a number of advantages when tuned to maximal twist: (i) in this case automatic $\mathcal{O}(a)$ improvement is obtained by tuning only one parameter, the bare untwisted quark mass, while avoiding additional tuning of operator-specific improvement-coefficients; (ii) the mixing pattern in the renormalisation process can be significantly simplified; (iii) the twisted mass provides an infra-red regulator helping to overcome possible problems with ergodicity in molecular dynamics based algorithms.

The main drawback of the twisted mass approach is the explicit breaking of parity and isospin symmetry which are only restored when the continuum limit is reached. However, due to $\mathcal{O}(a)$ improvement, this breaking is an $\mathcal{O}(a^2)$ effect as confirmed by simulations performed in the quenched approximation[13,14]. Furthermore, theoretical considerations indicate that these $\mathcal{O}(a^2)$ effects may become large *only* for the particular case of the neutral pion mass[15]. This expectation is confirmed by results from numerical simulations for other quantities considered so far.

The condition of tuning the parameters of the theory such that a parity odd operator vanishes can be reformulated as the condition that the so-called PCAC mass is tuned to zero. Here the PCAC mass

$$m_{\rm PCAC} = \frac{\sum_{\mathbf{x}} \langle \partial_0 A_0^a(x) P^a(0) \rangle}{2 \sum_{\mathbf{x}} \langle P^a(x) P^a(0) \rangle}, \qquad a = 1, 2 \tag{3}$$

is evaluated at large enough time separation, such that the pion ground state is dominant.

The strategy we have followed is to take the value of $am_{\rm crit}$ from the simulation at the lowest available value $a\mu_{\rm min} \ll a\Lambda_{\rm QCD}$. In this situation $\mathcal{O}(a)$ improvement is still guaranteed, because working at $\mu_{\rm min}$ merely leads to $\mathcal{O}(a\mu_{\rm min}\Lambda_{\rm QCD})$ effects in $m_{\rm crit}$ and $\mathcal{O}(a^2\mu_{\rm min}\Lambda_{\rm QCD})$ relative corrections in physical quantities[9]. Although these theoretical arguments show that we are left with only $\mathrm{O}(a^2)$ lattice artefacts, numerical computations are required to determine the coefficient multiplying the a^2-term for the observables of interest. In our simulations the condition of maximal twist has been numerically realised with good accuracy, which in this context means $m_{\rm PCAC}(\mu_{\rm min})/\mu_{\rm min} < a\Lambda_{QCD}$ within statistical errors ($a\Lambda_{QCD} \sim 0.1$ in our case).

The algorithm we used is a HMC algorithm with mass preconditioning[16,17] and multiple time scale integration described in detail in Ref.[18]. The trajectory length τ was set to $\tau = 1/2$ in all our runs. While the plaquette autocorrelation times are typically O(10-50), for quantities such as $am_{\rm PS}$ or $af_{\rm PS}$ we find them to be substantially smaller, typically by a factor of 5-10.

2.2 $f_{\rm PS}$ and $m_{\rm PS}$ As a Function of the Quark Mass

The *charged* pseudo scalar meson mass $am_{\rm PS}$ is as usual extracted from the time exponential decay of appropriate correlation functions. In contrast to pure Wilson fermions, for maximally twisted mass fermions an exact lattice Ward identity allows to extract the (charged) pseudo scalar meson decay constant $f_{\rm PS}$ with no need to compute any renormalisation constant since $Z_P = 1/{Z_\mu}^2$.

In our χPT based analysis, we take into account finite size corrections because on our lattices at the lowest and next-to-lowest μ-values they turn out to affect $am_{\rm PS}$ and, in particular, $af_{\rm PS}$ in a significant way. We have used continuum χPT to describe consistently the dependence of the data both on the finite spatial size (L) and on μ.

We fit the appropriate ($N_f = 2$) χPT formulae[19,20]

$$m^2_{\rm PS}(L) = 2B_0\mu \left[1 + \frac{1}{2}\xi\tilde{g}_1(\lambda)\right]^2 \left[1 + \xi\log(2B_0\mu/\Lambda_3^2)\right] , \tag{4}$$

$$f_{\rm PS}(L) = F \left[1 - 2\xi\tilde{g}_1(\lambda)\right] \left[1 - 2\xi\log(2B_0\mu/\Lambda_4^2)\right] , \tag{5}$$

to our raw data for $m_{\rm PS}$ and $f_{\rm PS}$ simultaneously. Here

$$\xi = 2B_0\mu/(4\pi F)^2 , \qquad \lambda = \sqrt{2B_0\mu L^2} . \tag{6}$$

The finite size correction function $\tilde{g}_1(\lambda)$ was first computed by Gasser and Leutwyler in Ref.[19] and is also discussed in Ref.[20] from which we take our notation (except that our normalisation of f_π is 130.7 MeV). In Eqs. (4) and (5) NNLO χPT corrections are assumed to be negligible. The formulae above depend on four unknown parameters, B_0, F, Λ_3 and Λ_4, which will be determined by the fit. For $\mu = 0.004$ and $\mu = 0.0064$ we found the effect of finite size corrections to be 0.5% and 0.2% for the pseudo scalar mass and 2.2% and 0.9% for the pseudo scalar decay constant, respectively. For our larger values of μ the finite size corrections are negligible.

We determine $a\mu_\pi$, the value of $a\mu$ at which the pion assumes its physical mass, by requiring that the ratio $[\sqrt{[m^2_{\rm PS}(L=\infty)]}/f_{\rm PS}(L=\infty)]$ takes the value $(139.6/130.7) = 1.068$. From the knowledge of $a\mu_\pi$ we can evaluate $\bar{l}_{3,4} \equiv \log(\Lambda^2_{3,4}/m^2_\pi)$ and using f_π the value of the lattice spacing a in fm.

For our lightest four μ-values, we find an excellent fit to our data on $f_{\rm PS}$ and $m_{\rm PS}$ (see figures 1 and 2). The fitted values of the four parameters are

$$\begin{aligned} 2aB_0 &= 4.99(6) , \\ aF &= 0.0534(6) , \\ \log(a^2\Lambda_3^2) &= -1.93(10) , \\ \log(a^2\Lambda_4^2) &= -1.06(4) . \end{aligned} \tag{7}$$

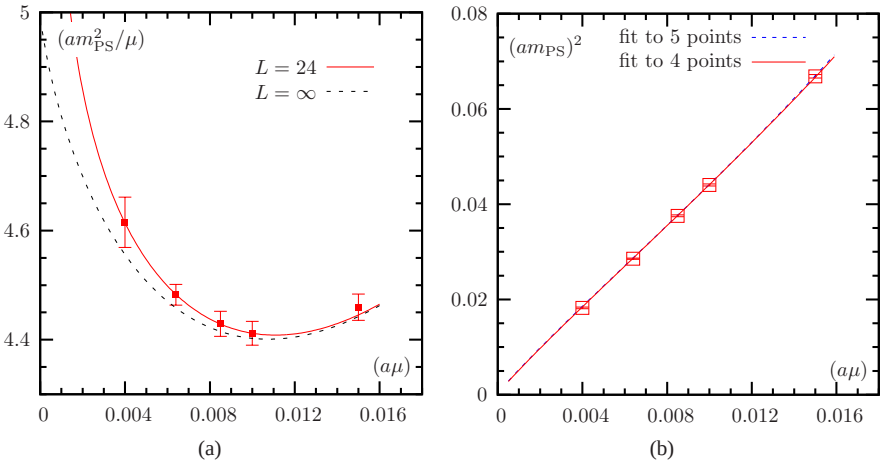


Figure 1. In (a) we show $(am_{\rm PS})^2/(a\mu)$ as a function of $a\mu$. We plot the χPT fit of Eq. (4) applied to the raw data on the $L = 24$ lattice from the lowest four μ-values. We represent the finite size correction by the dashed line. In (b) we show $(am_{\rm PS})^2$ as a function of $a\mu$. Here we present two χPT fits with Eq. (4), one taking all data points and one leaving out the point at the largest value $a\mu = 0.015$. Also in figure (b) we show the $L = 24$ data points.

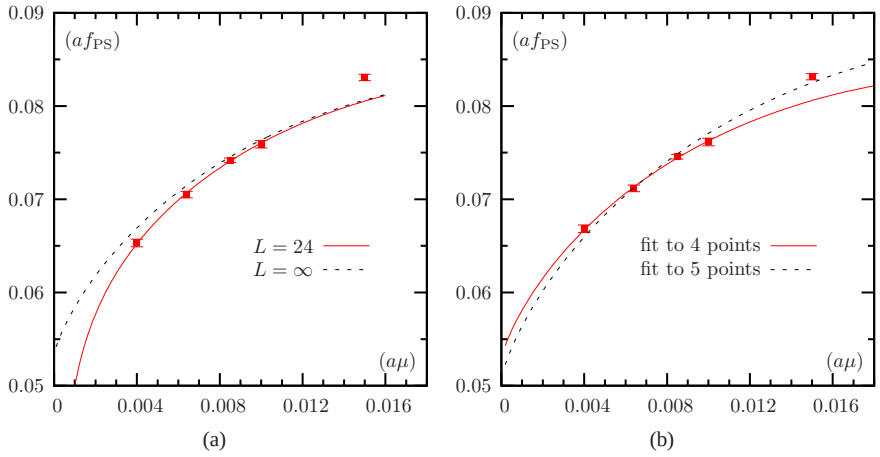


Figure 2. We show $af_{\rm PS}$ as a function of $a\mu$ together with fits to χPT formula Eq. (5). In (a) we show the fit applied to the raw data on the $L = 24$ lattice at the 4 lowest values of $a\mu$. We represent the finite size correction by the dashed curve. In (b) we present two fits, one taking all data and one leaving out the point at the largest value $a\mu = 0.015$. Here we show only the finite size corrected ($L \to \infty$) data points.

Our data are clearly sensitive to Λ_3 as visualised in figure 1(a). We obtain

$$a\mu_\pi = 0.00078(2), \qquad \bar{l}_3 = 3.65(12), \qquad \bar{l}_4 = 4.52(06) \tag{8}$$

which compares nicely with other determinations (for a review see Ref.[21]).

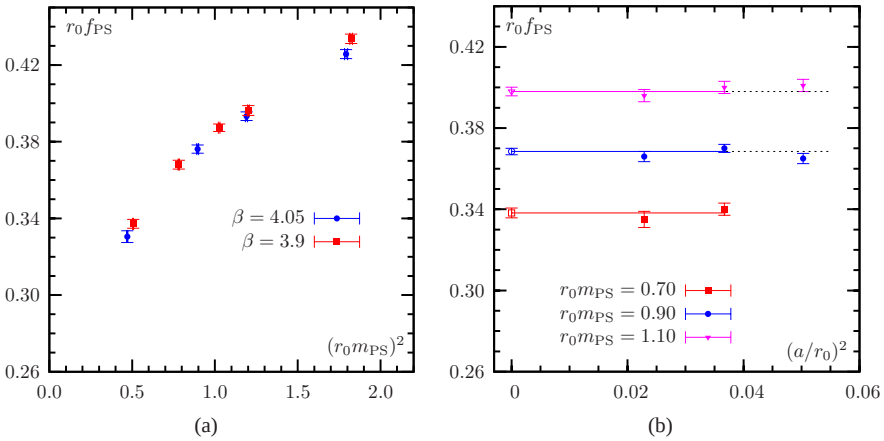


Figure 3. (a) $r_0 f_{\rm PS}$ as a function of $(r_0 m_{\rm PS})^2$ for $\beta = 3.9$ and $\beta = 4.05$. (b) Continuum extrapolation of $f_{\rm PS}$ at fixed volume for three reference values of $r_0 m_{\rm PS}$. The data points at $\beta = 3.8$ are not used.

3 Summary

In this contribution we have presented results of simulations of lattice QCD with $N_f = 2$ maximally twisted Wilson quarks. We reached a pseudo scalar meson mass of about $300\ \mathrm{MeV}$. The numerical stability and smoothness of the simulations allowed us to obtain precise results for the pseudo scalar mass and decay constant which in turn led to determine the low energy constants of the effective chiral Lagrangian. In particular, we find for the pseudo scalar decay constant in the chiral limit $F = 121.3(7)\ \mathrm{MeV}$, and $\bar{l}_3 = 3.65(12)$ and $\bar{l}_4 = 4.52(6)$ where only statistical errors are given. Our results at the two values of the lattice spacing which are used for a continuum extrapolation show almost negligible lattice spacing artefacts which can be seen from fig. 3, see also refs.[7,22], where we show the results for the pseudo scalar decay constant. In the left panel, we exhibit $r_0 f_{\rm PS}$ for two values of the lattice spacing demonstrating that the data scale very nicely. In the right panel we show $r_0 f_{\rm PS}$ at fixed values of the pseudo scalar mass, $r_0 m_{\rm PS}$. In this figure we also show results for an additional value of the lattice spacing which is, however, not used for a continuum extrapolation since it is presently not clear whether our physical condition to reach maximal twist is realized here.

Besides the here described determination of the low energy constants, our collaboration is also computing many more physical observables from the existing configurations from maximally twisted mass fermions. These are the full octet and decuplet spectra[23], moments of parton distribution functions[24], non-perturbatively obtained renormalization constants using the RI-MON scheme[25], light quark masses and decay constants[26,27], charm physics[28], meson form factors[29], exploration of the ϵ-regime of chiral perturbation theory[30], study of cut-off effects at tree-level of perturbation theory[31], neutral mesons[32], overlap fermions on twisted mass sea quarks[33,34] and twisted mass fermions at non-vanishing temperature[35].

All these results obtained from maximally twisted mass fermions and two mass degenerated flavours of quarks are very encouraging and promising. They suggest that indeed

maximally twisted mass fermions provide a very valuable formulation of lattice QCD. It is therefore most natural to extend the present study by incorporating the strange and the charm quarks as dynamical degrees in the simulation. This is possible using the setup of ref.[36] and has already been explored in ref.[37]. The simulations for this realistic setup of QCD are presently ongoing and the tuning to maximal twist is already well advanced. Thus, it is to be expected that similar precise results for many physical quantities can be obtained in the near future.

Acknowledgments

An essential amount of computer time for this project was made available to us by the John von Neumann-Institute for Computing on the JUMP and JUBL systems. I am most grateful to all my collaborators of the ETMC. Without their hard work and enthusiasm the nice and precise results presented here would not have been possible to obtain.

References

1. K. G. Wilson. Confinement of quarks. *Phys. Rev.*, D10:2445–2459, 1974.
2. R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz. Lattice QCD with a chirally twisted mass term. *JHEP*, 08:058, 2001.
3. R. Frezzotti and G. C. Rossi. Chirally improving Wilson fermions. I: O(a) improvement. *JHEP*, 08:007, 2004.
4. K. Jansen. Actions for dynamical fermion simulations: Are we ready to go? *Nucl. Phys. Proc. Suppl.*, 129:3–16, 2004.
5. Ph. Boucaud et al. Dynamical twisted mass fermions with light quarks. *Phys. Lett.*, B650:304–311, 2007.
6. A. Shindler. Twisted mass lattice qcd. 2007.
7. Carsten Urbach. Lattice qcd with two light wilson quarks and maximally twisted mass. 2007.
8. P. Weisz. Continuum limit improved lattice action for pure Yang-Mills theory. 1. *Nucl. Phys.*, B212:1, 1983.
9. R. Frezzotti, G. Martinelli, M. Papinutto, and G. C. Rossi. Reducing cutoff effects in maximally twisted lattice QCD close to the chiral limit. *JHEP*, 04:038, 2006.
10. Andrea Shindler. Twisted mass lattice QCD: Recent developments and results. *PoS*, LAT2005:014, 2006.
11. F. Farchioni et al. Exploring the phase structure of lattice QCD with twisted mass quarks. *Nucl. Phys. Proc. Suppl.*, 140:240–245, 2005.
12. F. Farchioni et al. The phase structure of lattice QCD with Wilson quarks and renormalization group improved gluons. *Eur. Phys. J.*, C42:73–87, 2005.
13. K. Jansen et al. Flavour breaking effects of Wilson twisted mass fermions. *Phys. Lett.*, B624:334–341, 2005.
14. F. Farchioni et al. Twisted mass fermions: Neutral pion masses from disconnected contributions. *PoS*, LAT2005:033, 2006.
15. Roberto Frezzotti and Giancarlo Rossi. O(a^2) cutoff effects in wilson fermion simulations. 2007.

16. M. Hasenbusch. Speeding up the Hybrid-Monte-Carlo algorithm for dynamical fermions. *Phys. Lett.*, B519:177–182, 2001.
17. M. Hasenbusch and K. Jansen. Speeding up lattice QCD simulations with clover-improved Wilson fermions. *Nucl. Phys.*, B659:299–320, 2003.
18. C. Urbach, K. Jansen, A. Shindler, and U. Wenger. HMC algorithm with multiple time scale integration and mass preconditioning. *Comput. Phys. Commun.*, 174:87–98, 2006.
19. J. Gasser and H. Leutwyler. Light quarks at low temperatures. *Phys. Lett.*, B184:83, 1987.
20. Gilberto Colangelo, Stephan Dürr, and Christoph Haefeli. Finite volume effects for meson masses and decay constants. *Nucl. Phys.*, B721:136–174, 2005.
21. H. Leutwyler. pi pi scattering. 2006.
22. Petros Dimopoulos, Roberto Frezzotti, Gregorio Herdoiza, Carsten Urbach, and Urs Wenger. Scaling and low energy constants in lattice qcd with nf=2 maximally twisted wilson quarks. 2007.
23. : Constantia Alexandrou et al. Baryon masses with dynamical twisted mass fermions. 2007.
24. Remi Baron et al. Moments of meson distribution functions with dynamical twisted mass fermions. 2007.
25. P. Dimopoulos et al. Renormalisation of quark bilinears with nf=2 wilson fermions and tree-level improved gauge action. 2007.
26. Vittorio Lubicz, Silvano Simula, and Cecilia Tarantino. Light quark masses and pseudoscalar decay constants from nf=2 twisted mass qcd. 2007.
27. B. Blossier et al. Light quark masses and pseudoscalar decay constants from nf=2 lattice qcd with twisted mass fermions. 2007.
28. B. Blossier, G. Herdoiza, and S. Simula. Twisted mass qcd in the charm sector. 2007.
29. S. Simula. Pseudo-scalar meson form factors with maximally twisted wilson fermions at nf = 2. 2007.
30. K. Jansen, A. Nube, A. Shindler, C. Urbach, and U. Wenger. Exploring the epsilon regime with twisted mass fermions. 2007.
31. K. Cichy, J. Gonzalez Lopez, K. Jansen, A. Kujawa, and A. Shindler. Cutoff effects for wilson twisted mass fermions at tree- level of perturbation theory. 2007.
32. Chris Michael and Carsten Urbach. Neutral mesons and disconnected diagrams in twisted mass qcd. 2007.
33. O. Bär, K. Jansen, S. Schaefer, L. Scorzato, and A. Shindler. Overlap fermions on a twisted mass sea. 2006.
34. Nicolas Garron and Luigi Scorzato. Results from overlap valence quarks on a twisted mass sea. 2007.
35. Ernst-Michael Ilgenfritz et al. Twisted mass qcd at finite temperature. 2007.
36. R. Frezzotti and G. C. Rossi. Chirally improving Wilson fermions. II: Four-quark operators. *JHEP*, 10:070, 2004.
37. T. Chiarappa et al. Numerical simulation of QCD with u, d, s and c quarks in the twisted-mass Wilson formulation. 2006.

Hadron Spectrum of QCD with one Quark Flavour

Federico Farchioni[1], **István Montvay**[2], **Gernot Münster**[1], **Enno E. Scholz**[3], **Tobias Sudmann**[1], **and Jaïr Wuilloud**[1]

[1] Universität Münster, Institut für Theoretische Physik
Wilhelm-Klemm-Strasse 9, D-48149 Münster, Germany
E-mail: munsteg@uni-muenster.de

[2] Deutsches Elektronen-Synchrotron DESY
Notkestr. 85, D-22603 Hamburg, Germany

[3] Physics Department, Brookhaven National Laboratory
Upton, NY 11973, USA

The hadron spectrum of Quantum Chromodynamics with a single flavour of quarks is studied by Monte Carlo simulations at two different lattice spacings. The lightest simulated quark mass corresponds to a pion with mass ~ 270 MeV. The symmetry of the single flavour theory can be artificially enhanced by adding extra valence quarks, which can be interpreted as u and d quarks. Properties of the theory are analyzed by making use of the ideas of partially quenched chiral perturbation theory applied to the extended theory.

1 Introduction and Motivation

For the strong interactions of elementary particles a successful theoretical description exists in terms of Quantum Chromodynamics (QCD). In the low-energy regime of QCD perturbation theory ceases to be applicable. In particular, to calculate the spectrum of physical particles from QCD is a genuine non-perturbative problem. This is where numerical simulations of QCD by means of the Monte Carlo method turn out to be an indispensable and very powerful tool.

To include the dynamics of quarks in numerical simulations is an important but demanding task. It requires the calculation of the determinant of the quadratic form describing fermions in the QCD action, and consumes most of the computer time. Due to the development of computer power, it has become possible in recent years to include the dynamics of quarks in QCD simulations.

Quarks do not exist as isolated particles, but bind together in mesons and baryons. Therefore it is not possible to give a straightforward unique definition of quark masses. In recent years M. Creutz has drawn the attention to open problems in QCD, which show up in this context[1,2]. One question raised by Creutz, having a relevant phenomenological impact, is whether it is possible to define in an unambiguous way the case where *one* quark (say the u quark) becomes massless. The arguments against an unique definition of the massless limit[1] especially hold for the one-flavour theory.

A difference between QCD with more than one quark flavour ($N_f > 1$) and QCD with a single flavour ($N_f = 1$), which plays a crucial role in this context, is the nature of chiral symmetry. The Lagrangian of QCD with $N_f > 1$ flavours of massless quarks possesses a continuous chiral symmetry with symmetry group $\mathrm{SU}(N_f)_L \otimes \mathrm{SU}(N_f)_R \otimes \mathrm{U}(1)_A$. In the quantum theory the axial $\mathrm{U}(1)_A$ is not a symmetry due to the Adler-Bell-Jackiw anomaly.

The remaining $\mathrm{SU}(N_f)_L \otimes \mathrm{SU}(N_f)_R$ chiral symmetry is broken spontaneously to its diagonal subgroup, which is the usual $\mathrm{SU}(N_f)$ vector symmetry. The Goldstone bosons associated with the spontaneous symmetry breakdown are massless pseudoscalar mesons, e.g. the pions and kaons. If quark masses m_q are turned on, the Goldstone bosons acquire masses. If the quark masses are equal and small, the masses of the pseudo-Goldstone bosons are approximately given by the Gell-Mann-Oakes-Renner relation

$$m_{PS}^2 = 2Bm_q\,, \tag{1}$$

where B is a constant. In such a situation the point, where the pseudo-Goldstone bosons become massless, uniquely marks the vanishing of the quark masses.

QCD with one flavour of quarks ($N_f = 1$ QCD) radically differs from QCD with two or more flavours due to the absence of a chiral symmetry: there is no $\mathrm{SU}(N_f)_L \otimes \mathrm{SU}(N_f)_R$ symmetry group, and the abelian $\mathrm{U}(1)_A$ symmetry of the one-flavour theory is washed out at the quantum level by the anomaly. As a consequence of this, the main features of the phase structure and mass spectrum of the single flavour theory strongly deviate from the familiar picture. In particular, there are no pseudo-Goldstone bosons, which would indicate the point of vanishing quark masses. In our project we undertake a thorough investigation of $N_f = 1$ QCD in order to study these fundamental questions.

A second interesting aspect is the possibility of a spontaneous breaking of CP (charge conjugation and parity) symmetry in QCD for special choices of the quark masses, conjectured for the first time by Dashen[3]. According to the Vafa-Witten theorem[4], a prerequisite for the spontaneous breaking of a discrete symmetry is a non-positive fermion measure, which in $N_f = 1$ QCD is possible for negative quark masses. The transition line is indeed expected to be located[5] on the negative real quark mass axis in the extended complex parameter space.

Another intriguing aspect of one-flavour QCD is the connection with the $\mathcal{N}$=1 supersymmetric Yang-Mills theory (SYM) in a particular limit of a large number of colours N_c[6]. An important place where relics of SUSY in $N_f = 1$ QCD can be investigated, is the low-lying bound-state spectrum[7]. Low-energy models for SYM[8] predict a low-lying chiral supermultiplet including two scalar particles with opposite parity. In $N_f = 1$ QCD these two particles can be identified with the η and the σ meson. Their mass ratio including $O(1/N_c)$ corrections is expected to be $m_\sigma/m_\eta = N_c/(N_c-2)$[9].

2 Partially Quenched QCD

As we have argued in[7], it is useful to enhance the symmetry of the one-flavour theory by adding extra valence quarks which are *quenched*, i.e., not included in the fermion determinants. We take two valence quarks u and d with masses m_V and one sea quark s with mass m_S. For our purpose the case of degenerate valence and sea quark mass $m_V = m_S$ is particularly convenient. In this case the combined sea and valence sector is characterized by an exact $\mathrm{SU}(N_F)$ flavour symmetry, where $N_F \equiv N_V + N_f = 2 + 1$.

At the point of vanishing quark masses (see below) the generic partially quenched theory has a graded symmetry, which is broken spontaneously into a "flavour" symmetry $\mathrm{SU}(N_F|N_V)$, also valid for non-vanishing degenerate quark masses. The $\mathrm{SU}(N_F)$ subgroup represents the flavour symmetry in the combined sea and valence quark sectors. It

implies that the hadronic bound states appear in exactly degenerate SU(N_F) multiplets for $m_V = m_S$.

In particular, this extended theory contains a degenerate octet of pseudoscalar mesons ("pions" π^a, $a = 1, \ldots, 8$) satisfying an $\mathrm{SU}(3)$-symmetric PCAC relation. From the divergence of the axial-vector current and pseudoscalar density a bare *PCAC quark mass* can be naturally defined. Due to the exact SU(3)-symmetry, the corresponding renormalized quark mass $m^{\mathrm{R}}_{\mathrm{PCAC}}$ can be defined by an SU(3)-symmetric multiplicative renormalization.

As we will confirm numerically in Sec. 5, the masses of the "pions" can be made to vanish by suitably tuning the bare quark mass on the lattice. In this situation the renormalized quark mass vanishes, too. We stress here that the pions are not particles in the physical spectrum of the theory. Nevertheless, their properties as mass and decay constant are well defined quantities which can be computed on the lattice. The same applies for the PCAC quark mass, which can therefore be regarded as a potential candidate for a definition of the quark mass in this theory.

2.1 Chiral Perturbation Theory

We have calculated both the masses and decay constants of the pseudo-Goldstone bosons[7] in next-to-leading order of partially quenched chiral perturbation theory along the lines of Ref. 12, including $\mathcal{O}(a)$ lattice artifacts[11]. Leaving out the lattice corrections, the expressions for the renormalized pion masses and decay constants in terms of the (renormalized) PCAC quark mass are

$$\begin{aligned} m_\pi^2 &= \chi_{\mathrm{PCAC}} + \frac{\chi^2_{\mathrm{PCAC}}}{16\pi^2 F_0^2} \ln \frac{\chi_{\mathrm{PCAC}}}{\Lambda_3^2}\,, \\ \frac{f_\pi^{\mathrm{R}}}{F_0\sqrt{2}} &= 1 - \frac{\chi_{\mathrm{PCAC}}}{32\pi^2 F_0^2} \ln \frac{\chi_{\mathrm{PCAC}}}{\Lambda_4^2}\,, \end{aligned} \tag{2}$$

where $\chi_{\mathrm{PCAC}} = 2B_0\, m^{\mathrm{R}}_{\mathrm{PCAC}}$ with the usual leading order low-energy constant B_0, and Λ_3 and Λ_4 are next-to-leading order low-energy constants.

The pion mass and the mass of the "physical" η can be related by an extension of PQChPT[10,12]. The leading order expression for the mass of the η reads

$$m_\eta^2 = \frac{m_\Phi^2 + \chi_{\mathrm{PCAC}}}{1+\alpha}, \tag{3}$$

where α and m_Φ are free parameters in this context. Our numerical results for m_η allow to determine α and m_Φ (see Section 5).

3 Simulation

The present study[7] has been performed on a $12^3 \cdot 24$ lattice at inverse gauge coupling $\beta = 3.8$ and $16^3 \cdot 32$ lattice at $\beta = 4.0$. The lattice spacings correspond to $a \simeq 0.19\,\mathrm{fm}$ and $a \simeq 0.13\,\mathrm{fm}$, respectively. We use the Sommer parameter[13] r_0 for setting the scale, fixed at the conventional value $r_0 \equiv 0.5\,\mathrm{fm}$. The extensions of the lattices are roughly constant: $L = 2.23\,\mathrm{fm}$ and $L = 2.14\,\mathrm{fm}$.

For the fermions we use the Wilson action; in the SU(3) gauge sector we apply the tree-level improved Symanzik (tlSym) action[14]. The update algorithm is a Polynomial Hybrid

Monte Carlo algorithm (PHMC)[15,16] allowing the simulation of an odd number of fermion species. The present version[17] is based on a two-step polynomial approximation of the inverse fermion matrix with stochastic correction in the update chain, taken over from the two-step multi-boson algorithm of Ref. 18. We reached a relatively high total acceptance of 0.64–0.72. A correction factor in the measurement is associated with configurations for which eigenvalues of the (squared Hermitian) fermion matrix $Q^2[U]$ lie outside the validity interval of the polynomial approximation.

As indicated in the introduction, the sign of the quark determinant is an important issue in $N_f = 1$ QCD. With Wilson lattice fermions for small quark masses, it can become negative due to quantum fluctuations. The sign $\sigma[U]$ of the fermion determinant $\det Q[U]$ has to be included in the reweighting of the configurations. For its computation we applied two methods. In the first we studied the *spectral flow* of the Hermitian fermion matrix[19], following Ref. 20 for the computation of the low-lying eigenvalues. Alternatively, we computed the spectrum of the non-Hermitian matrix by ARPACK Arnoldi routines[21], concentrating on the lowest real eigenvalues: sign changes are signaled by negative real eigenvalues.

In most of our simulations the quark mass is large enough to prevent sign changes and the occurrence of a negative determinant is a rare event. On the other hand, for the smallest quark masses, $m_q \simeq 12\,\mathrm{MeV}$, a sizeable effect can be observed: the sign of the determinant pushes up the hadron masses by $7 - 10\%$.

4 Hadron Spectrum

For the meson states we consider the simplest interpolating operators in the pseudoscalar and scalar sectors:

$$\eta(0^-): \quad P(x) = \bar{\psi}(x)\gamma_5\psi(x)\,, \tag{4}$$

$$\sigma(0^+): \quad S(x) = \bar{\psi}(x)\psi(x)\,. \tag{5}$$

Corresponding states in the QCD spectrum are the $\eta'(958)$ and $f_0(600)$ (or σ). The disconnected diagrams of the hadron correlators of η and σ were computed by applying stochastic sources with complex Z_2 noise and spin dilution[22].

In the baryon sector we considered a spin $3/2$ operator constructed from the quark field. The associated low lying state is expected to be the positive parity $(\frac{3}{2}^+)$ one. This corresponds to the $\Delta^{++}(1232)$ of QCD if our dynamical fermion is interpreted as a u quark.

We used the single spatial plaquette to obtain the mass of the 0^{++} glueball state, employing APE smearing[23] and variational methods[24] to obtain optimal operators. Nevertheless, our statistics turned out not to be large enough to obtain an accurate estimate of the glueball masses.

The results for the hadron masses are plotted in Fig. 1 as a function of the bare PCAC quark mass. Since we use physical units here, results from the two lattice spacings can be compared. The scaling is satisfactory for the case of η, whose mass could be computed with the best accuracy. The determination of the σ meson mass seems to require large statistics.

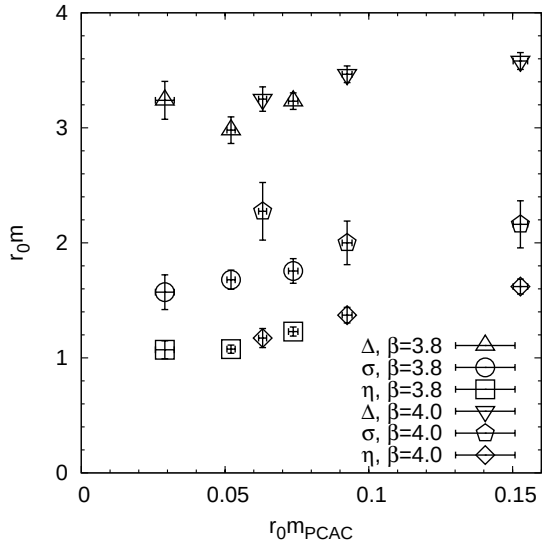


Figure 1. The mass of the lightest physical particles in one-flavour QCD as a function of the bare PCAC quark mass. The masses are multiplied by the scale parameter r_0 in order to obtain dimensionless quantities.

5 Partially Quenched Analysis

We have obtained results for the pion masses and decay constant and for the nucleon mass in the partially quenched sector. The partially quenched ChPT formulae are used to extract the corresponding low-energy coefficients from the pion data. Considering the number of lattice data at our disposal, a fit using the continuum terms and neglecting lattice artifacts is possible. We fitted the data for both β values simultaneously neglecting the dependence of the renormalization factors Z_A and Z_P upon the lattice coupling constant. The data and the fitted curves are shown in Fig. 2. In order to improve the numerical results for the universal low-energy constants $\Lambda_{3,4}$, which do not explicitly depend on the lattice spacing a, we also performed fits to ratios of these quantities for the data at $\beta = 4.0$. We obtain in this case the results

$$\frac{\Lambda_3}{F_0} = 10.0 \pm 2.6\,, \tag{6}$$

$$\frac{\Lambda_4}{F_0} = 31.5 \pm 14.3\,, \tag{7}$$

which, interestingly, are compatible with phenomenological values obtained from ordinary QCD[25]. The errors are, however, quite large and we hope to improve these determinations in the future.

In addition, we investigated the relation between the mass of the pion and of the physical η by fitting simultaneously m_π^2 and m_η^2 as a function of the PCAC quark mass according to formula (3), again considering only $\beta = 4.0$. Our data are consistent with a vanishing

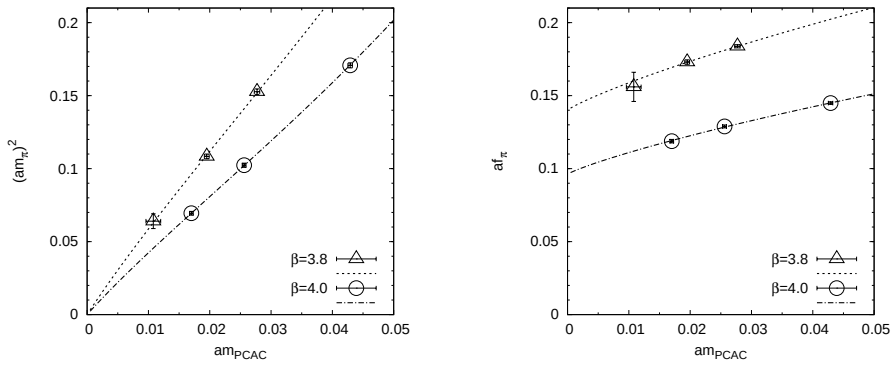


Figure 2. Pion masses squared and pion decay constants in lattice units and the results of the PQChPT fit.

α. Setting $\alpha = 0$, we find

$$m_\Phi = 284 \pm 40\,\mathrm{MeV} \tag{8}$$

in physical units. The value of m_Φ can also be obtained from the Witten-Veneziano formula[26]

$$m_\Phi^2 = \frac{4N_f}{(f_\pi^{\mathrm{R}})^2}\chi_t \tag{9}$$

valid at leading-order in the ('t Hooft) large N_c limit. An estimate of the quenched topological susceptibility[27] is $\chi_t = (193 \pm 9\,\mathrm{MeV})^4$. Using our value for f_π^R, which is subject to a sizeable statistical error, one would obtain $m_\Phi = 450 \pm 170$ MeV.

6 Summary and Outlook

This first Monte Carlo investigation of $N_f = 1$ QCD reveals the qualitative features of the low lying hadron spectrum of this theory. The lightest hadron is the pseudoscalar η meson while the scalar meson, the σ, is about a factor 1.5 heavier. It is interesting to compare our data with the estimate in[28] $m_\sigma/m_\eta \simeq N_c/(N_c-2) = 3$ for $N_c = 3$. The above prediction applies for the massless theory and one could expect the agreement to improve for smaller quark masses. Our bare quark masses (estimated from the PCAC quark mass in the valence analysis) range between 10 MeV and 60 MeV, while the lightest pion mass is ~ 270 MeV.

The lightest baryon, the Δ ($\frac{3}{2}^+$), is by about a factor 3 heavier than the η meson. The lightest scalar mass obtained with a glueball 0^{++} operator lies between the σ meson and the Δ baryon mass. However, this mass could be overestimated, since, due to the high level of noise, only small time-separations could be included in the analysis.

In general, the mass measurements have relatively large errors between 3–10%. In order to obtain more quantitative results, larger statistics and smaller quark masses are required. We hope to be able to make progresses in both directions[29] with our new simulations using Stout-smeared links in the fermion action. Some preliminary results have already been obtained.

The introduction of a partially quenched extension of the single flavour theory with valence quarks allows to define the bare quark mass in terms of the PCAC quark mass of the fictitious multi-flavour theory. Comparison of lattice data with partially quenched chiral perturbation theory allowed the determination of some of the low-energy constants of the chiral Lagrangian. The latter are compatible, even if with large error, with recent lattice determinations for $N_f = 2$ QCD.

A further direction of investigation for the future[29] is the CP-violating phase transition expected at negative quark masses[5]. For this aspect of the single flavour theory the non-positivity of the fermion measure plays an essential role.

Acknowledgments

We are grateful to Luigi Scorzato for valuable discussions and help. We thank the John von Neumann Institute for Computing (NIC) and the computer center at DESY Hamburg for providing us the necessary technical help and computer resources. This work is supported in part by the Deutsche Forschungsgemeinschaft under grant Mu757/13-1. E.S. is supported by the U.S. Dept. of Energy under contract DE-AC02-98CH10886.

References

1. M. Creutz, *Ambiguities in the up-quark mass*, Phys. Rev. Lett. **92**, 162003, 2004; hep-ph/0312225.
2. M. Creutz, *Spontaneous violation of CP symmetry in the strong interactions*, Phys. Rev. Lett. **92**, 201601, 2004; hep-lat/0312018.
3. R.F. Dashen, *Some features of chiral symmetry breaking*, Phys. Rev. D **3**, 1879, 1971.
4. C. Vafa and E. Witten, *Restrictions on symmetry breaking in vector-like gauge theories*, Nucl. Phys. B **234**, 173, 1984.
5. M. Creutz, *Quark masses and chiral symmetry*, Phys. Rev. D **52**, 2951, 1995; hep-th/9505112.
6. A. Armoni, M. Shifman and G. Veneziano, *Exact results in non-supersymmetric large N orientifold field theories*, Nucl. Phys. B **667**, 170, 2003; hep-th/0302163; A. Armoni and M. Shifman, *The cosmological constant and domain walls in orientifold field theories and N = 1 gluodynamics*, Nucl. Phys. B **670**, 148, 2003; hep-th/0303109; A. Armoni, M. Shifman and G. Veneziano, *SUSY relics in one-flavor QCD from a new 1/N expansion*, Phys. Rev. Lett. **91**, 191601, 2003; hep-th/0307097.
7. F. Farchioni, I. Montvay, G. Münster, E. E. Scholz, T. Sudmann and J. Wuilloud, *Hadron masses in QCD with one quark flavour*, Eur. Phys. J. C **52**, 305, 2007; arXiv:0706.1131 [hep-lat].
8. G. Veneziano and S. Yankielowicz, *An effective Lagrangian for the pure N=1 supersymmetric Yang-Mills theory*, Phys. Lett. B **113**, 231, 1982.
9. A. Armoni and E. Imeroni, *Predictions for orientifold field theories from type 0' string theory*, Phys. Lett. B **631**, 192, 2005; hep-th/0508107.
10. C. W. Bernard and M. F. L. Golterman, *Partially quenched gauge theories and an application to staggered fermions*, Phys. Rev. D **49**, 486, 1994; hep-lat/9306005.
11. G. Rupak and N. Shoresh, *Chiral perturbation theory for the Wilson lattice action*, Phys. Rev. D **66**, 054503, 2002; hep-lat/0201019.

12. S. R. Sharpe and N. Shoresh, *Partially quenched chiral perturbation theory without* Φ_0, Phys. Rev. D **64**, 114510, 2001; hep-lat/0108003.
13. R. Sommer, *A new way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory*, Nucl. Phys. B **411**, 839, 1994; hep-lat/9310022.
14. P. Weisz, *Continuum limit improved lattice action for pure Yang-Mills theory. 1*, Nucl. Phys. B **212**, 1, 1983; P. Weisz and R. Wohlert, *Continuum limit improved lattice action for pure Yang-Mills theory. 2*, Nucl. Phys. B **236**, 397, 1984 [Erratum-ibid. B **247**, 544, 1984].
15. P. de Forcrand and T. Takaishi, *Fast fermion Monte Carlo*, Nucl. Phys. Proc. Suppl. **53**, 968, 1997; hep-lat/9608093.
16. R. Frezzotti and K. Jansen, *A polynomial hybrid Monte Carlo algorithm*, Phys. Lett. B **402**, 328, 1997; hep-lat/9702016.
17. I. Montvay and E. Scholz, *Updating algorithms with multi-step stochastic correction*, Phys. Lett. B **623**, 73, 2005; hep-lat/0506006; E. E. Scholz and I. Montvay, *Multi-step stochastic correction in dynamical fermion updating algorithms*, PoS **LAT2006**, 037, 2006; hep-lat/0609042.
18. I. Montvay, *An algorithm for gluinos on the lattice*, Nucl. Phys. B **466**, 259, 1996; hep-lat/9510042.
19. R. G. Edwards, U. M. Heller and R. Narayanan, *Spectral flow, chiral condensate and topology in lattice QCD*, Nucl. Phys. B **535**, 403, 1998; hep-lat/9802016.
20. T. Kalkreuter and H. Simma, *An accelerated conjugate gradient algorithm to compute low lying eigenvalues: A study for the Dirac operator in SU(2) lattice QCD*, Comput. Phys. Commun. **93**, 33, 1996; hep-lat/9507023.
21. R. B. Lehoucq, D. C. Sorensen and C. Yang, 1997, `http://www.caam.rice.edu/software/ARPACK/`.
22. F. Farchioni and R. Peetz, *The low-lying mass spectrum of the N = 1 SU(2) SUSY Yang-Mills theory with Wilson fermions*, Eur. Phys. J. C **39**, 87, 2005; hep-lat/0407036.
23. M. Albanese *et al.* [APE Collaboration], *Glueball masses and string tension in lattice QCD*, Phys. Lett. B **192**, 163, 1987.
24. C. J. Morningstar and M. J. Peardon, *The glueball spectrum from an anisotropic lattice study*, Phys. Rev. D **60**, 034509, 1999; hep-lat/9901004.
25. J. Gasser and H. Leutwyler, *Chiral perturbation theory to one loop*, Annals Phys. **158**, 142, 1984; S. Dürr, *M(pi)**2 versus m(q): Comparing CP-PACS and UKQCD data to chiral perturbation theory*, Eur. Phys. J. C **29**, 383, 2003; hep-lat/0208051.
26. E. Witten, *Current algebra theorems for the U(1) Goldstone boson*, Nucl. Phys. B **156**, 269, 1979; G. Veneziano, *U(1) Without Instantons*, Nucl. Phys. B **159**, 213, 1979.
27. S. Dürr, Z. Fodor, C. Hoelbling and T. Kurth, *Precision study of the SU(3) topological susceptibility in the continuum*, JHEP **0704**, 055, 2007; hep-lat/0612021.
28. F. Sannino and M. Shifman, *Effective Lagrangians for orientifold theories*, Phys. Rev. D **69**, 125004, 2004; hep-th/0309252.
29. F. Farchioni, I. Montvay, G. Münster, E. E. Scholz, T. Sudmann and J. Wuilloud, in preparation.

The Transverse Spin Structure of the Nucleon

G. Schierholz

Deutsches Elektronen-Synchrotron DESY
E-mail: Gerrit.Schierholz@desy.de

– For the *QCDSF Collaboration* –

Recent advances in algorithms and computer technology have enabled lattice QCD to make increasingly quantitative calculations of hadron structure. In this contribution we shall present first results on the transverse spin structure of nucleon and pion from simulations of $N_f = 2$ flavours of light dynamical sea quarks.

1 Introduction

A major scientific goal of elementary particle physics is to achieve a quantitative understanding of the structure and interactions of hadrons in terms of their quark and gluon constituents. Although a quarter of a century has passed since the experimental discovery of quarks in the nucleon and the invention of QCD, understanding how QCD works remains one of the great puzzles in physics.

Generalized Parton Distributions (GPDs)[1,2] are the modern tool to deliver a detailed description of the microscopic structure of hadrons. They allow to map out the longitudinal and transverse distributions of quarks and gluons in the fast moving hadron, thus providing an essentially holographic picture of mesons and baryons. This is illustrated in Fig. 1,

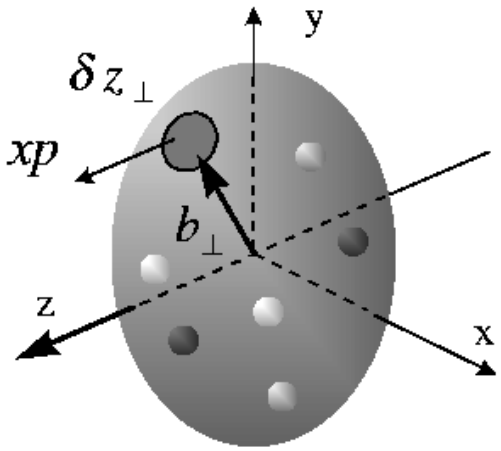


Figure 1. Pictorial view of the nucleon moving with momentum p in the z-direction in terms of its quark constituents.

where we show

$$\rho^q(x, b_\perp) = H^q(x, b_\perp^2)\,, \tag{1}$$

which describes the probability of finding an unpolarized quark of type q with fractional longitudinal momentum x at impact parameter $b_\perp$ in the unpolarized nucleon. The impact parameter $b_\perp$ denotes the distance between the quark and the centre of momentum of the nucleon in the plane perpendicular to its momentum. In different kinematic limits of the GPDs one recovers the familiar elastic form factors, parton densities and distribution amplitudes. For example,

$$q(x) = \int \frac{d^2 b_\perp}{(2\pi)^2} H^q(x, b_\perp^2)\,, \tag{2}$$

where $q(x)$ is the ordinary quark parton distribution of the nucleon. GPDs are directly relevant to present and future experiments at CERN, DESY, GSI, JLab, MAMI and RHIC-spin.

Recent advances in lattice field theory, numerical algorithms and computer technology have enabled us to compute GPDs from first principles. In this contribution I shall report first results on the spin structure of nucleon and pion.

2 Spin Structure of the Nucleon

Let us consider a fast moving nucleon[a] with transverse spin $S_\perp$ (with respect to the nucleon's momentum)[3]. We are interested in the density of quarks $\rho^q(x, b_\perp, s_\perp, S_\perp)$, $q = u$ and d, with fractional momentum x and transverse spin $s_\perp$. In lattice calculations only moments of $\rho(x, b_\perp, s_\perp, S_\perp)$ are accessible[4]:

$$\begin{aligned}\rho^{q\,n}(b_\perp, s_\perp, S_\perp) = \int_{-1}^{1} dx\, x^{n-1} \rho^q(x, b_\perp, s_\perp, S_\perp) = \frac{1}{2}\Big\{ A_n^q(b_\perp^2) + s_\perp^i S_\perp^i \Big(A_{Tn}^q(b_\perp^2) \\ - \frac{1}{4m_N^2} \Delta_{b_\perp} \widetilde{A}_{Tn}^q(b_\perp^2) \Big) + \frac{b_\perp^j \epsilon^{ji}}{m_N} \Big(S_\perp^i B_n^{q\,\prime}(b_\perp^2) + s_\perp^i \overline{B}_{Tn}^{q\,\prime}(b_\perp^2) \Big) \\ + s_\perp^i (2 b_\perp^i b_\perp^j - b_\perp^2 \delta^{ij}) S_\perp^j \frac{1}{m_N^2} \widetilde{A}_{Tn}^{q\,\prime\prime}(b_\perp^2) \Big\}\,,\end{aligned} \tag{3}$$

where m_N is the nucleon mass and A_n, A_{Tn}, $\widetilde{A}_{Tn}$, B_n and B_{Tn} are generalized form factors (GFFs). The momentum space GFFs are obtained from off-forward nucleon matrix elements of composite operators. For example,

$$\begin{aligned}\langle p', S' |\, \mathcal{O}_T^{q\,\mu\nu}\, | p, S \rangle = \overline{u}(p', S') \Big\{ \sigma^{\mu\nu} \gamma_5 \big(A_{T1}^q(t) - \frac{t}{2m_N^2} \widetilde{A}_{T1}^q(t) \big) \\ + \frac{\epsilon^{\mu\nu\alpha\beta} \Delta_\alpha \gamma_\beta}{2m_N} \overline{B}_{T1}^q(t) - \frac{\Delta^{[\mu} \sigma^{\nu]\alpha} \gamma_5 \Delta_\alpha}{2m_N^2} \widetilde{A}_{T1}^q(t) \Big\} u(p, S)\,,\end{aligned} \tag{4}$$

[a] Note that the parton distributions are boost independent. However, the variable x maps into the momentum fraction of quarks and gluons only in the formalism of light-cone quantization, which is equivalent to Feynman's parton model in the infinite momentum frame $p_z \to \infty$. In the rest frame of the nucleon, for example, x is just a special combination of the parton's off-shell energy k_0 and momentum k_z.

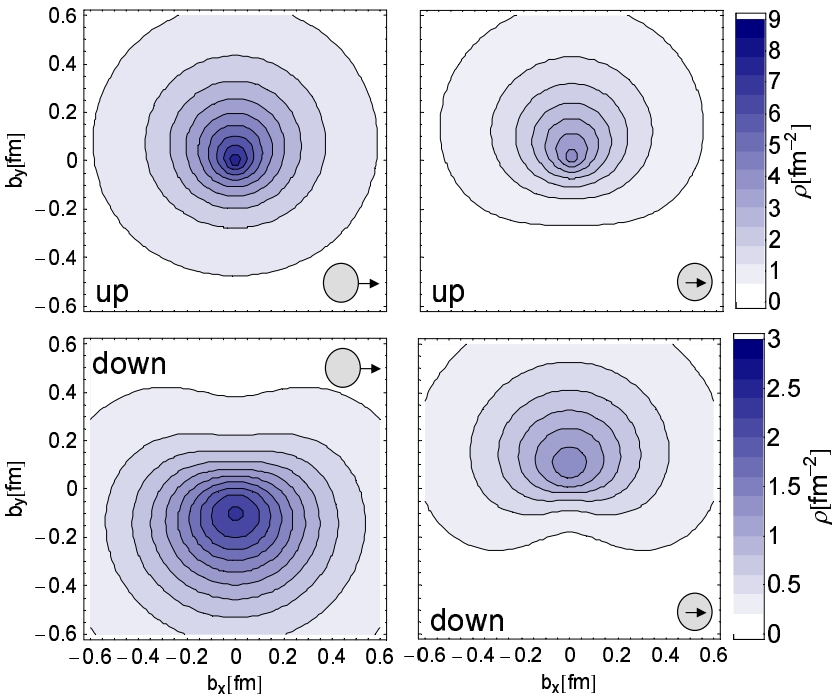


Figure 2. The first moment $\rho^{q\,1}(b_\perp, s_\perp, S_\perp)$ for unpolarized quarks in the transversely polarized proton (left panels) and transversely polarized quarks in the unpolarized nucleon (right panels) separately for up (top panels) and down (bottom panels) quarks. The direction of the quark (nucleon) spin is indicated by the inner (outer) arrows.

where $\mathcal{O}_T^{q\,\mu\nu} = \bar{q}\sigma^{\mu\nu}\gamma_5 q$ is the lowest operator of a tower of quark helicity flip operators, and $t = (p' - p)^2$. The impact parameter space GFFs that enter in (3) are obtained by Fourier transform

$$X(b_\perp^2) = \int \frac{d^2\Delta_\perp}{(2\pi)^2} \mathrm{e}^{-ib_\perp \Delta_\perp} X(t = -\Delta_\perp^2)\,, \tag{5}$$

where $\Delta_\perp$ is the transverse momentum transfer to the nucleon.

The calculations are done on gauge field configurations with $N_f = 2$ flavours of light dynamical sea quarks. The calculation proceeds in four steps. In the first step we compute the GFFs on any of our lattices, which differ essentially by the quark mass and the lattice spacing a. In the second step we fit the lattice data to a suitable parameterization. It turns out that the data can be well described by a p-pole ansatz $X(t) = X(0)/(1 - t/(pm_X^2))^p$. In the third step we extrapolate the results to the physical quark mass. To do so, we make use of the predictions of chiral perturbation theory (ChPT) as far as possible. The interested reader will find a more detailed discussion of the general procedure in the literature[5–7]. Finally we Fourier transform the GFFs to impact parameter space. We do not see any significant dependence on the lattice spacing a in the range of lattice spacings $0.07 \lesssim a \lesssim 0.1$ fm we have explored, indicating that cut-off effects are small. The

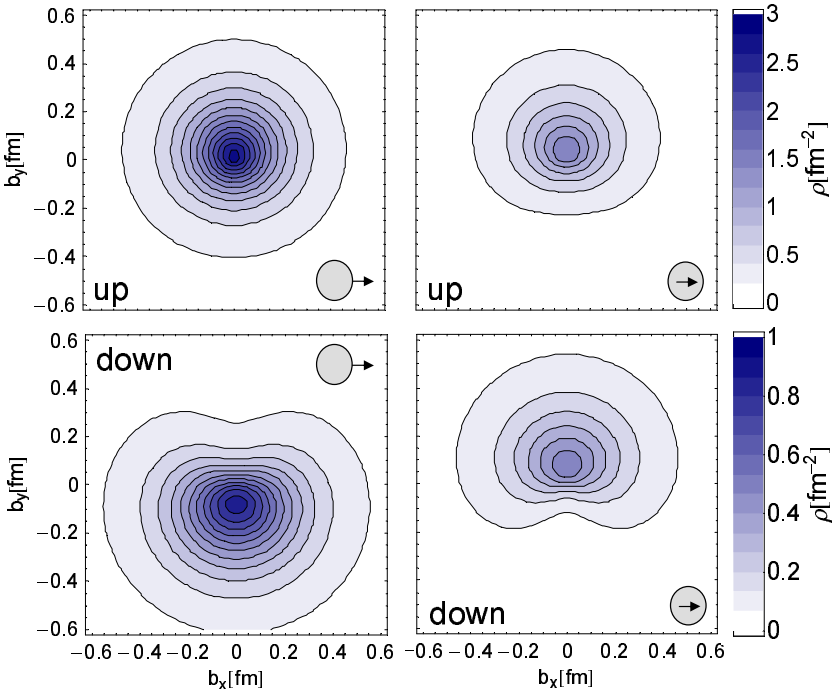


Figure 3. The second moment $\rho^{q\,2}(b_\perp, s_\perp, S_\perp)$. The notation and arrangement is the same as in Fig. 2.

hadronic matrix elements that enter the calculation of the GFFs need to be renormalized, which is done nonperturbatively[8].

In Fig. 2 we show the first moment $\rho^{q\,1}(b_\perp, s_\perp, S_\perp)$ separately for up and down quarks in the proton. We find strong distortions of unpolarized quarks in the transversely polarized proton (left panels), with the up and down quarks shifted in opposite direction. This distortion can be explained by orbital motion of the quarks[9], assuming that the orbital angular momentum L_q of the up (down) quark is negative (positive). Recent lattice calculations have shown that this is indeed the case. Note that there is no gluon transversity, which could mix with quarks under evolution. As a result, the transverse spin structure is essentially valence-like, and we expect that the transverse spin and transverse orbital angular momentum are simply aligned. For transversely polarized quarks in the unpolarized proton (right panels) the densities for up and down quarks are both shifted in the positive y-direction. This is due to rather large and positive values of the tensor GFFs $\overline{B}^u_{Tn}(t=0)$ and $\overline{B}^d_{Tn}(t=0)$. In Fig. 3 we show the second moment $\rho^{q\,2}(b_\perp, s_\perp, S_\perp)$. The pattern is very similar to that in Fig. 2. The main difference is that the spin densities are more peaked towards $b_\perp = 0$.

Further information about the spin structure of the nucleon is obtained from considering transversely polarized quarks in the transversely polarized nucleon. In Fig. 4 we plot the first moment $\rho^{q\,1}(b_\perp, s_\perp, S_\perp)$ for $S_\perp$ and $s_\perp$ being (anti-)parallel resp. orthogonal to each other. A rich pattern of distortions is observed.

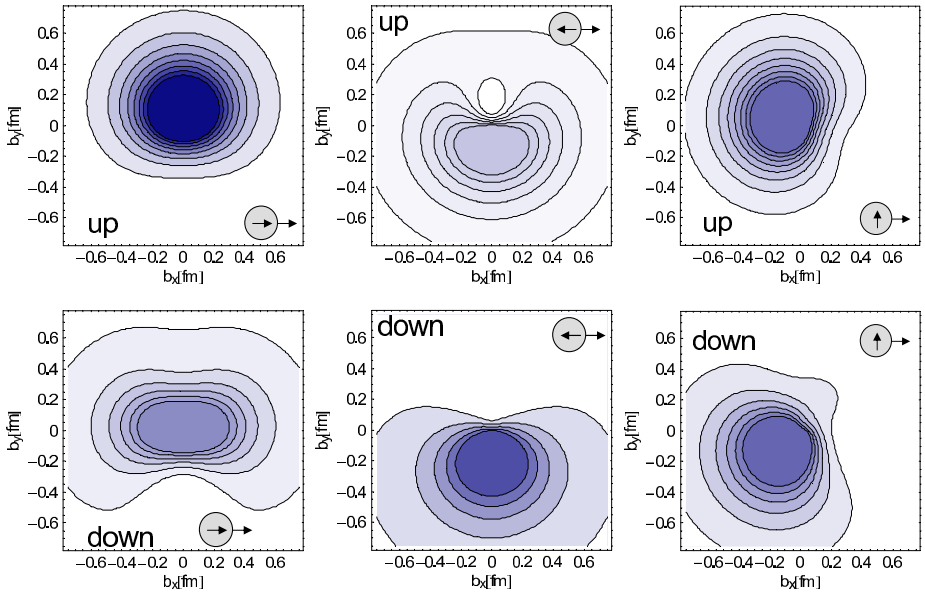


Figure 4. The first moment $\rho^{q\,1}(b_\perp, s_\perp, S_\perp)$ for nucleon and quarks both being polarized. The notation is the same as in Fig. 2.

3 Spin Structure of the Pion

Let us now turn to the pion[10]. As the pion has no spin, the impact parameter GPDs are much simpler than in the case of the nucleon. While the longitudinal spin structure of the pion is trivial, nothing is known about the transverse spin distribution. In this case the spin density reduces to

$$\rho^{q\,n}(b_\perp, s_\perp) = \int_{-1}^{1} dx\, x^{n-1} \rho^q(x, b_\perp, s_\perp) = \frac{1}{2}\Big\{ A_n^q(b_\perp^2) - \frac{b_\perp^j \epsilon^{ji} s_\perp^i}{m_\pi} B_{Tn}^{q\,\prime}(b_\perp^2) \Big\}\,, \quad (6)$$

and the GFFs B_{Tn}^q are obtained from off-forward pion matrix elements of the form

$$\langle p' |\, \mathcal{O}_T^{q\,\mu\nu} \,| p \rangle = \frac{\bar{p}^\mu \Delta^\nu - \Delta^\mu \bar{p}^\nu}{m_\pi}\, B_{T1}^q(t)\,, \quad (7)$$

where $\bar{p} = (p' + p)/2$.

In Fig. 5 we plot the first two moments of $\rho^q(x, b_\perp, s_\perp)$. Like in the nucleon, the transversely polarized quarks are substantially distorted, which reveals, for the first time, a nontrivial spin structure of the pion.

The fact, that the distortion in both the pion and nucleon is of the same sign and magnitude, suggests that $B_{Tn}^q(b_\perp^2)$ in the pion and $\overline{B}_{Tn}^q(b_\perp^2)$ in the nucleon are about equal. Both GFFs are directly related to the correlation between the quark's transverse spin and intrinsic transverse momentum through the Boer-Mulders function[11]. A possible explanation of this observation is that the quarks in the fast moving hadron are in an s-wave state with an admixture of p-wave by means of the lower Dirac component of the wave function[12].

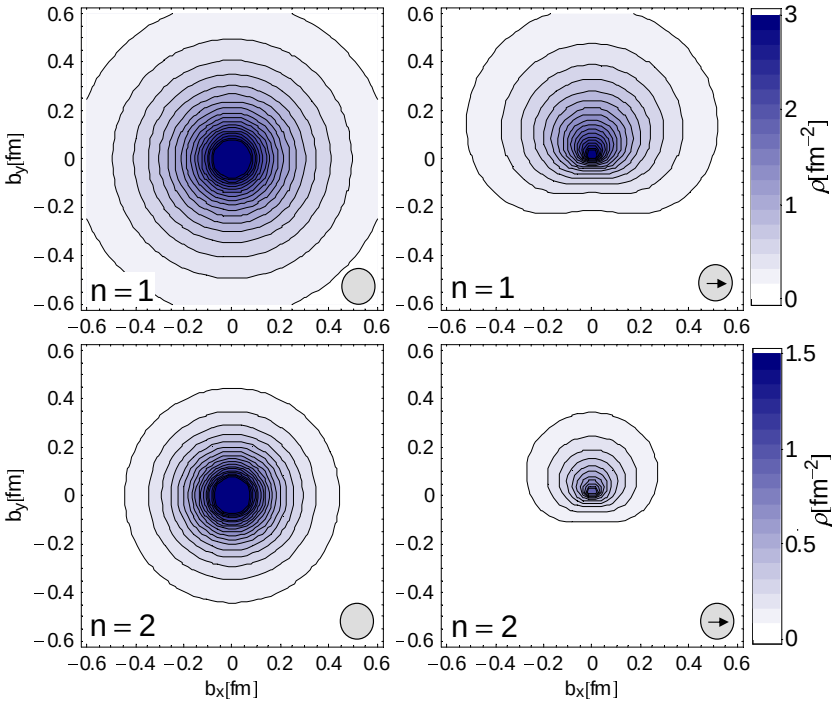


Figure 5. The first and second moment, $\rho^{q\,1}(b_\perp, s_\perp)$ and $\rho^{q\,2}(b_\perp, s_\perp)$, for unpolarized (left panels) and transversely polarized (right panels) up quarks in π^+.

As in case of the nucleon, we find that the second moment of $\rho^q(x, b_\perp, s_\perp)$ is more sharply peaked towards $b_\perp = 0$ than the first moment.

4 Conclusions

GPDs are a powerful tool in unravelling the microscopic structure of hadrons. We have seen that a profound understanding can be achieved through elaborate lattice simulations.

In the future we will need to extend our calculations to smaller quark masses and larger lattices to better constrain the parameterization of the GFFs as well as the extrapolation of the lattice results to the physical pion mass. We furthermore plan to extend the calculations to $N_f = 2 + 1$, and eventually $N_f = 2 + 1 + 1$, flavours of dynamical sea quarks.

Acknowledgments

The simulations of the dynamical gauge field configurations have been performed on the BlueGene/L at KEK (Tsukuba), on the BlueGene/L at EPCC (Edinburgh) and on the BlueGene/L at NIC (Jülich). The analysis of the gauge field configurations has been done on the ape1000 and apeNEXT computers at DESY (Zeuthen). We thank all institutions for

their support. We furthermore acknowledge financial support by the Deutsche Forschungsgemeinschaft DFG through FOR-465 (Forschergruppe Gitter-Hadronen-Phänomenologie) and by the European Union through I3HP (Contract No. RII3-CT-2004-506078).

References

1. M. Diehl, Phys. Rept. **388** (2003) 41 [hep-ph/0307382].
2. A. V. Belitsky and A. V. Radyushkin, Phys. Rept. **418** (2005) 1 [arXiv:hep-ph/0504030].
3. M. Göckeler, Ph. Hägler, R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow, A. Schäfer, G. Schierholz, H. Stüben and J. M. Zanotti, Phys. Rev. Lett. **98** (2007) 222001 [arXiv:hep-lat/0612032].
4. M. Diehl and Ph. Hägler, Eur. Phys. J. C **44** (2005) 87 [hep-ph/0504175].
5. M. Göckeler, Ph. Hägler, R. Horsley, Y. Nakamura, M. Ohtani, D. Pleiter, P. E. .L. Rakow, A. Schäfer, G. Schierholz, W. Schroers, H. Stüben and J. M. Zanotti, arXiv:0709.3370 [hep-lat].
6. D. Brömmel, M. Göckeler, Ph. Hägler, R. Horsley, Y. Nakamura, M. Ohtani, D. Pleiter, P. E. .L. Rakow, A. Schäfer, G. Schierholz, W. Schroers, H. Stüben and J. M. Zanotti, arXiv:0710.1534 [hep-lat].
7. M. Göckeler, Ph. Hägler, R. Horsley, Y. Nakamura, M. Ohtani, D. Pleiter, P. E. .L. Rakow, A. Schäfer, G. Schierholz, W. Schroers, H. Stüben and J. M. Zanotti, arXiv:0710.2159 [hep-lat].
8. G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa and A. Vladikas, Nucl. Phys. B **445** (1995) 81 [hep-lat/9411010]; M. Göckeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, P. E. L. Rakow, A. Schäfer, G. Schierholz and A. Schillerg, Nucl. Phys. B **544** (1999) 699 [hep-lat/9807044].
9. M. Burkardt, Phys. Rev. D **66** (2002) 114005 [arXiv:hep-ph/0209179].
10. D. Brömmel, M. Diehl, M. Göckeler, Ph. Hägler, R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow, A. Schäfer, G. Schierholz, H. Stüben and J. M. Zanotti, arXiv:0708.2249 [hep-lat].
11. D. Boer and P. J. Mulders, Phys. Rev. D **57** (1998) 5780 [arXiv:hep-ph/9711485].
12. M. Burkardt and B. Hannafious, arXiv:0705.1573 [hep-ph].

Exploring the $\Delta I = 1/2$ Rule in Non-Leptonic Kaon Decays

Pilar Hernández[1], Mikko Laine[2], Bastian Knippschild[3], Carlos Pena[4], Thomas Tilli[3], Emma Torró[1], Jan Wennekers[5], and Hartmut Wittig[3]

[1] Depto. de Física Teòrica and IFIC, Universitat de València, E-46100 Burjassot, Spain

[2] Faculty of Physics, University of Bielefeld, D-33501 Bielefeld, Germany

[3] Institut für Kernphysik, University of Mainz, D-55099 Mainz, Germany

[4] Departamento de Física Teórica C-XI and Instituto de Física Teórica UAM/CSIC C-XVI Universidad Autónoma de Madrid, Cantoblanco E-28049 Madrid, Spain

[5] School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK

A satisfactory theoretical explanation for the large enhancement of the $\Delta I = 1/2$ amplitude in $K \to \pi\pi$ decays has been lacking for more than 30 years. While it has long been suspected that the charm quark, and the fact that its mass is much larger than typical hadronic scales, could be the main source of the effect, it is only now that this scenario can be tested reliably via lattice simulations of QCD. In this note we give an account of our ongoing project to investigate the mechanism behind the $\Delta I = 1/2$ rule. So far we have found indirect evidence that a significant contribution to the enhancement must indeed come from the charm quark.

1 Introduction

The decays and mixing patterns of K-mesons, mediated by the weak interaction, are among the prime sources of information on discrete symmetries in Nature. In particular, the violation of CP symmetry, which transforms particles into anti-particles, has wide-ranging implications, since it may explain why matter dominates over anti-matter in the universe. In the Standard Model, CP violation is incorporated via a complex phase in the Cabibbo-Kobayashi-Maskawa matrix, and its value can be pinned down by studying the mixing between a neutral kaon, K^0, and its anti-particle, $\bar{K}^0$. Such a process gives rise to what is called *indirect* CP violation, while CP symmetry is *directly* violated in certain non-leptonic kaon decays, such as $K \to \pi\pi$.

Despite many efforts, our theoretical understanding of non-leptonic kaon decays on a quantitative basis is still rather limited. So far, attempts to compute the amount of direct versus indirect CP violation, which is parameterised by the ratio ϵ'/ϵ have not produced credible results. A closely related problem is the failure to explain the so-called $\Delta I = 1/2$ rule: if a neutral kaon, having isospin 1/2, decays into a pair of pions, the latter can either have isospin $I = 0$ or 2. The corresponding transition amplitudes are then given by the amplitudes A_0 and A_2 (up to a phase factor). The experimentally observed decay rates yield an unexpectedly large ratio of

$$A_0/A_2 \approx 22.1, \tag{1}$$

which implies that the decay in which isospin changes by 1/2 is favoured over the $\Delta I = 3/2$ transition by a large margin, and this observation is usually called the $\Delta I = 1/2$ rule.

Quantum Chromodynamics (QCD), the gauge theory of the strong interaction, should in principle allow to explain the large enhancement of the amplitude A_0. However, theoretical calculations based on perturbative QCD in conjunction with naïve estimates of the hadronic matrix elements involved, yield a value for A_0/A_2 that turns out to be smaller by a full order of magnitude. Thus, a successful explanation of the $\Delta I = 1/2$ rule must inevitably be based on a non-perturbative treatment of QCD. Numerical simulations of QCD on a space-time lattice are designed for this task. Nevertheless, the directly computation of amplitudes for non-leptonic decays such as $K \to \pi\pi$ remains a great challenge.

A cornerstone of our computational strategy is the matching of QCD to an effective low-energy description, called Chiral Perturbation Theory (ChPT). Since the amplitudes A_0, A_2 can be expressed in terms of the leading-order low-energy constants (LECs), g_1^+ and g_1^-, of the interaction Hamiltonian which describes weak non-leptonic kaon decays in ChPT, the main idea is to compute the LECs in lattice simulations. This can be achieved by comparing the predictions of ChPT for suitably chosen two- and three-point correlation functions in a finite volume with the analogous ones computed in lattice QCD for small masses and momenta.

In this note we present a progress report of our project [1–6] to investigate the mechanism of the $\Delta I = 1/2$ rule. In particular we seek to clarify whether the observed large enhancement in A_0 over A_2 has a single origin or is the result of an accumulation of several moderately large effects. To this end we specifically concentrate on the rôle of the charm quark. The fact that the mass of the latter of around 1.3 GeV is much greater than typical low-energy scales of a few hundred MeV, has led to the conjecture that its decoupling may produce a large enhancement of A_0, due to the appearance of additional operators with large hadronic matrix elements.[7] In our strategy we keep the charm quark as an "active" degree of freedom, which – in contrast to most other studies – is not integrated out from the theory.

In the first part of the project the LECs g_1^+, g_1^-, and, in turn, the amplitudes A_0 and A_2 are determined for the unphysical situation where the charm is degenerate with the light quark, $m_c = m_u = m_d = m_s$. In a second step we envisage monitoring the amplitudes for heavier charm, i.e. $m_c > m_u = m_d = m_s$.

2 Computational Strategy

The main ingredients of our strategy are described in [1,3] and are briefly summarised here. In the first stage of the calculation we focus on the case of a light and degenerate charm quark. Decays such as $K \to \pi\pi$ in which strangeness changes by one unit are described in QCD via the interaction Hamiltonian

$$\mathcal{H}_{\rm w} = \sqrt{2} G_{\rm F} (V_{us})^* V_{ud} \left\{ k_1^+ \mathcal{Q}_1^+ + k_1^- \mathcal{Q}_1^- \right\}, \tag{2}$$

where $G_{\rm F}$ is the Fermi constant and V_{us}, V_{ud} denote elements of the quark mixing matrix. The operators $\mathcal{Q}_1^\pm$ are expressed in terms of quark fields according to

$$\mathcal{Q}_1^\pm = \Big\{ (\overline{s}\gamma_\mu P_- u)(\overline{u}\gamma_\mu P_- d) \pm (\overline{s}\gamma_\mu P_- d)(\overline{u}\gamma_\mu P_- u) \Big\} - (u \to c). \tag{3}$$

The Wilson coefficients $k_1^\pm$ in the above expression absorb short-distance effects and can be computed reliably in perturbation theory.

In the chiral effective theory (ChPT), the corresponding Hamiltonian reads

$$\mathcal{H}_{\rm w}^{\rm ChPT} = 2\sqrt{2}G_{\rm F}(V_{us})^*V_{ud}\left\{g_1^+[\widehat{\mathcal{O}}_1^+] + g_1^-[\widehat{\mathcal{O}}_1^-]\right\}, \tag{4}$$

where the operators $\widehat{\mathcal{O}}_1^+$ and $\widehat{\mathcal{O}}_1^-$ are now expressed in terms of Goldstone boson fields rather than quark degrees of freedom. The *a priori* unknown LECs $g_1^\pm$ incorporate the short-distance effects from the strong interaction in this effective low-energy description. In the chiral limit, they are directly related to the amplitudes A_0 and A_2 via

$$\frac{A_0}{A_2} = \frac{1}{\sqrt{2}}\left(\frac{1}{2} + \frac{3}{2}\frac{g_1^-}{g_1^+}\right). \tag{5}$$

Thus, the determination of the LECs in lattice simulations of QCD gives a handle to compute the ratio of amplitudes. The key quantities which allow to determine $g_1^\pm$ are correlation functions of the four-quark operators $\mathcal{Q}_1^\pm$ and the left-handed current, i.e.

$$C_1^\pm(x_0,y_0) = \sum_{\vec{x},\vec{y}}\left\langle(\overline{d}\gamma_0 P_- u)(x)\,\mathcal{Q}_1^\pm(0)\,(\overline{u}\gamma_0 P_- s)(y)\right\rangle, \quad P_- = \tfrac{1}{2}(1-\gamma_5). \tag{6}$$

We are particularly interested in ratios of correlation functions, such as

$$R^\pm(x_0,y_0) = \frac{C_1^\pm(x_0,y_0)}{C(x_0)C(y_0)}, \tag{7}$$

where the two-point function of the left-handed current is given by

$$C(x_0) = \sum_{\vec{x}}\left\langle(\overline{d}\gamma_0 P_- u)(x)\,(\overline{u}\gamma_0 P_- s)(0)\right\rangle. \tag{8}$$

With these definitions one may formulate a matching condition, which allows to express the unknown LECs in terms of the ratios $R^\pm$, the latter of which are computable in lattice simulations. At large Euclidean times x_0, y_0 the ratios $R^\pm$ are proportional to $g_1^\pm$:

$$k_1^\pm(Z^\pm/Z_{\rm A}^2)\,R^\pm = \left\{1 + K^\pm\right\}\,g_1^\pm, \tag{9}$$

where the $K^\pm$ parameterise chiral corrections. The latter manifest themselves in terms of a dependence of the ratios $R^\pm$ on the volume (i.e. the box size L) and the quark mass. Thus, by computing $R^\pm$ for a range of finite volumes and/or quark masses, knowledge of $K^\pm$ serves to extract the LECs via eq. (9). Two kinematical regimes of QCD are of particular importance in this regard: the so-called ϵ-regime [8] is characterised by considering the chiral limit in a finite volume, so that the pion's correlation length exceeds the box size L, and the latter is the only scale left in the theory. Moreover, correlation functions may depend strongly on the topological properties of the gauge field in this regime [9]. On the other hand, in the so-called p-regime the pions still fit in the box, and chiral corrections can be computed as a power series in $(m_\pi L)^2$.

In our strategy we combine numerical data for $R^\pm$ from the two kinematical regimes in order to have better control over the mass and volume dependence, which, in turn, should lead to a more reliable determination of the LECs. The expressions for the chiral corrections $K^\pm$ at next-to-leading order in the chiral expansion have been published in[1,10] and[4] for the ϵ- and p-regimes, respectively.

In addition to the chiral corrections, the matching condition, eq. (9), also contains short-distance corrections. These are given by the Wilson coefficients $k_1^\pm$, which can be computed reliably in perturbation theory, as well as the renormalisation factors $Z^\pm$ of the

four-quark operators and $Z_{\rm A}$ of the axial current. The latter must be considered in the matching condition to account for the proper renormalisation of the operators whose correlation functions appear in the ratios $R^{\pm}$. As part of our project we have determined the renormalisation factors non-perturbatively [5], by applying the technique originally proposed in [11].

3 Numerical Simulations and Results

The most important ingredient in our simulations is the use of a fermionic discretisation which preserves chiral symmetry at finite lattice spacing. As pointed out in [12], this is the case for any Dirac operator which satisfies the Ginsparg-Wilson relation.[13]

Preserving chiral symmetry at all stages of the calculation is indispensable to allow for a reliable matching to ChPT. However, it comes at a price, since the implementation of the Neuberger-Dirac operator is numerically very costly. This is further exacerbated in the chiral regime, which is susceptible to numerical instabilities as well as strong statistical fluctuations, which are generated by the appearance of arbitrarily small eigenvalues of the Dirac operator.

In our simulations we employ the Neuberger-Dirac operator.[14] We have developed several numerical techniques in order to control the effects of the small eigenmodes. They include the determination of the topological index ν via the counting of zero modes, in the course of which the latter can also be determined with sufficient numerical accuracy.[15] Second, one can speed up the numerical inversion of the Neuberger-Dirac operator considerably by applying "low-mode preconditioning".[15] Finally, the large statistical fluctuations encountered for small quark masses and, in particular, in the ϵ-regime can be tamed by applying "low-mode averaging".[16,17] Without going into detail here, we refer the reader to the original articles cited above.

Our simulations have been performed in the quenched approximation, such that the effects of dynamical quarks are unaccounted for. Despite the fact that one incurs an unknown systematic error, one may argue that the use of the quenched approximation is quite sufficient for the sake of investigating the origins of the $\Delta I = 1/2$ rule, since one tries to explain an enhancement of a decay amplitude by a full order of magnitude, while typical quenching effects amount to $10 - 15\%$. Our main results have been obtained on lattices of size $16^3{\cdot}32$ for a value of the bare coupling $\beta \equiv 6/g_0^2 = 5.8485$, which corresponds to a lattice spacing in physical units of $a \simeq 0.12\,\mathrm{fm}$. In the ϵ-regime we computed the ratios $R^{\pm}$ on 746 configurations for two values of the quark mass. Results were divided into bins of fixed topological charge $|\nu|$, and a weighted average over the results computed in each bin was taken inside an interval $2 \leq |\nu| \leq 10$. In the p-regime $R^{\pm}$ was determined on 197 configurations at four values of the quark mass without distinguishing configurations in different topological sectors. Plots of the (unrenormalised) ratios $R^{\pm}$ obtained in both kinematical regimes as a functions of the bare quark mass are shown in Fig. 1.

The fits shown in the figure yield estimates for the bare LECs. After including the Wilson coefficients and the renormalisation factors, we obtain the final results [3]

$$g_1^+ = 0.51(3)(5)(6), \qquad g_1^- = 2.6(1)(3)(3), \tag{10}$$

which apply in the case of a light degenerate charm quark considered in our study. The quoted errors arise from statistics, the matching to ChPT and from the uncertainty in the non-perturbative determination of the renormalisation factors. From these results one may

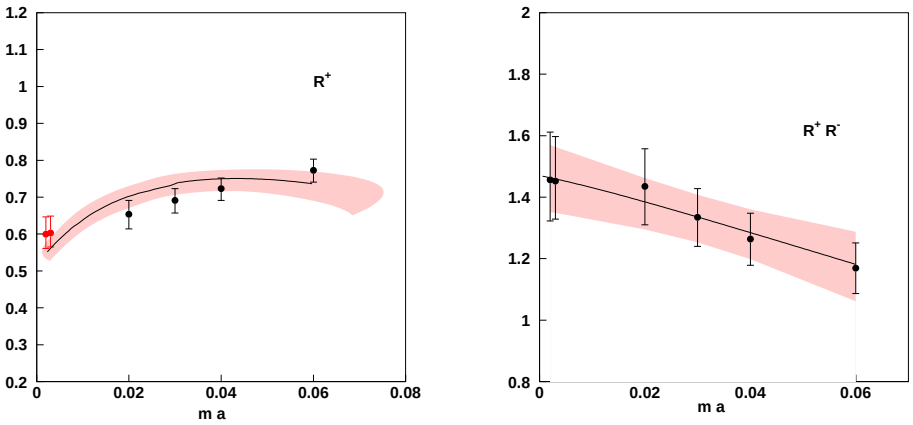


Figure 1. The ratio R^+ and the product R^+R^-. The two left-most points in each plot represent the data obtained in the ϵ-regime, while the remaining points lie in the p-regime. The shaded red bands denote the error from a joint fit of the data in both kinematical regimes to the expressions of ChPT.

compute the amplitudes A_0 and A_2 and compare to the physical result. First, one observes that $g_1^- \gg g_1^+$ so that our results indeed imply a clear hierarchy, producing an enhancement of A_0 over A_2 due to non-perturbative effects. Moreover, our findings indicate that the physical $\Delta I = 3/2$ amplitude A_2 is quite accurately reproduced by our result for g_1^+. However, the observed hierarchy is not sufficient to fully explain the physical enhancement of A_0, which is underestimated by a factor 4. Hence, our results support the conjecture that the decoupling of the charm quark plays a significant rôle for the $\Delta I = 1/2$ rule, despite the fact that our findings so far represent only indirect evidence for this scenario.

4 $g_1^\pm$ from Zero-Mode Wavefunctions

In the presence of gauge fields with non-trivial topology, $\nu \neq 0$, the (massless) Dirac operator exhibits $|\nu|$ zero modes of definite chirality, according to the Atiyah-Singer index theorem. This can be exploited for an alternative determination of $g_1^\pm$, by considering three-point functions of $\mathcal{Q}_1^\pm$ with the left-handed currents replaced by the corresponding pseudoscalar densities, i.e.

$$(\overline{u}\gamma_0 P_- s)(x) \to i(\overline{u}\gamma_5 s)(x), \qquad (\overline{d}\gamma_0 P_- u)(x) \to i(\overline{d}\gamma_5 u)(x). \tag{11}$$

In the ϵ-regime and for $\nu \neq 0$, the corresponding correlation functions may develop poles in $1/(mV)$, whenever the zero modes give a non-vanishing contribution. Since the residues of the poles admit a chiral expansion parameterised in terms of the same set of LECs, the alternative strategy is to compute the residues rather than the correlation functions themselves in the simulation.

The advantage of this strategy is that the zero modes can be obtained with relatively little numerical effort in the course of our calculation. Moreover, the systematic effects in this approach will be different from those of our standard method, thereby providing an independent cross-check on our findings described above.

To be more specific, we list a few more definitions. We consider the residue of the three-point function

$$\mathcal{C}^{\sigma}_{1;\,\nu} = \lim_{m\to 0}(mV)^2 \sum_{\vec{x},\vec{y}} \left\langle (\overline{d}\gamma_5 u)(x)\, \mathcal{Q}^{\sigma}_1(z)\, (\overline{u}\gamma_5 s)(y)\right\rangle_{\nu}$$
$$\equiv A_\nu(x_0-z_0, y_0-z_0) + \sigma \tilde{A}_\nu(x_0-z_0, y_0-z_0), \quad \sigma = \pm, \tag{12}$$

where A_ν and $\tilde{A}_\nu$ can be expressed in terms of zero-mode wavefunctions. For instance, if the zero modes have negative chirality, the expression for A_ν reads

$$A_\nu = \frac{1}{L^3}\sum_{\vec{x},\vec{y},\vec{z}} \left\langle \sum_{i\in\mathcal{K}} v_i^\dagger(x) S(x,z)\gamma_\mu P_- v_i(z) \sum_{j\in\mathcal{K}} v_j^\dagger(y) S(y,z)\gamma_\mu P_- v_j(z)\right\rangle_{\nu}, \tag{13}$$

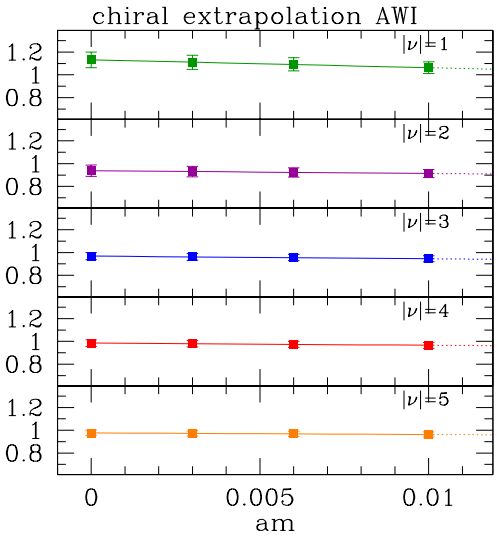


Figure 2. Verification of the chiral Ward identity, based on correlation functions saturated by zero modes, in various topological sectors.

where v_i is a zero mode. A similar expression can be derived for $\tilde{A}_\nu$. Equation (13) illustrates that a considerable amount of CPU time can be saved if one computes the above residues of correlation functions, in which some of the quark propagators are replaced by the zero mode contributions. In order to determine A_ν and $\tilde{A}_\nu$, the number of inversions of the Dirac operator is equal to twice the topological charge, i.e. $2|\nu|$, (with sources $v_i(x)$ and $v_i(y)$, since x_0 and y_0 need to be fixed). At the same time one can average over all spatial positions of the three sources, thereby reducing statistical fluctuations. Such averaging was only possible in the standard method of[1] through low-mode averaging (LMA), and only for the contribution of the low-modes. The price of LMA is $12 + 2 \times N_{\rm low}$ inversions, where $N_{\rm low}$ is the number of low modes treated separately. Typically $N_{\rm low}$ can be as large as 20, and hence the numerical cost of LMA is usually far greater than that required for the computation of the residues, where $|\nu|$ normally is not larger than 5.

In order to formulate the matching condition, we also consider the residues of the two-point functions of the pseudoscalar density and the left-handed axial current:

$$\mathcal{C}_\nu(x_0-z_0) \equiv \lim_{m\to 0}(mV)\sum_{\vec{x}} \left\langle (\overline{d}\gamma_5 u)(x)(\overline{u}\gamma_0 P_- d)(z)\right\rangle_{\nu}. \tag{14}$$

In practice it is advantageous to consider temporal derivatives of the correlators, and hence we define the ratio of correlation functions as

$$\mathcal{R}^{\pm}_\nu(x_0,y_0) = \frac{\partial_{x_0}\partial_{y_0}\mathcal{C}^{\pm}_{1;\,\nu}(x_0-z_0, y_0-z_0)}{\partial_{x_0}\mathcal{C}_\nu(x_0-z_0)\partial_{y_0}\mathcal{C}_\nu(y_0-z_0)}, \tag{15}$$

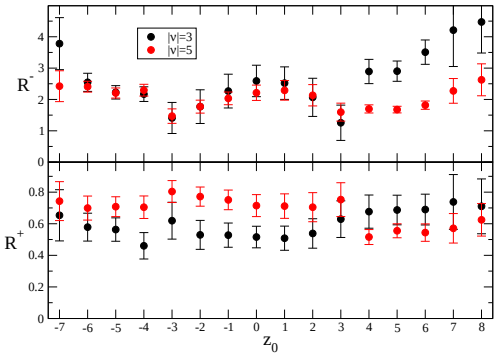


Figure 3. The ratios $\mathcal{R}_\nu^\pm$ for $|\nu| = 3$ and 5 computed on lattice of size 16^4 at $\beta = 5.8458$ for $x_0 = 5$ and $y_0 = 11$. The accumulated statistics amounts to only 27 and 24 configurations for $|\nu| = 3$ and 5, respectively.

while the matching condition now reads

$$k_1^\pm (Z^\pm/Z_{\rm A}^2)\, \mathcal{R}_\nu^\pm = \left\{1 + \mathcal{K}_\nu^\pm\right\}\, g_1^\pm . \tag{16}$$

The chiral corrections $\mathcal{K}_\nu^\pm$ have been worked out in ChPT at next-to-leading order.[6] It remains to compute the ratios $\mathcal{R}_\nu^\pm$ in simulations and extract the LECs $g_1^\pm$ through eq. (16).

An important consistency check of the procedure, which involves the calculation of two-point functions only, is the verification of the chiral Ward identity. In the chiral limit one expects

$$Z_{\rm A}\partial_{x_0}\mathcal{C}_\nu(x_0 - y_0) = \lim_{m\to 0}(m^2 V)\sum_{\vec{x}} \left\langle (\overline{d}\gamma_5 u)(x)(\overline{u}\gamma_5 d)(y)\right\rangle_\nu . \tag{17}$$

In practice, the above correlators are evaluated for small, but non-zero quark masses. After taking the renormalisation factor of the axial current, $Z_{\rm A}$, into account [18], the ratio of the two sides of eq. (17) is expected to extrapolate to unity, and indeed this is what we observe (see Fig. 2). The calculation of the residues of three-point functions on lattices of size 16^4 and 24^4 is currently under way. A plot of the preliminary data for the ratios $\mathcal{R}_\nu^\pm$ is shown in Fig. 3.

5 Conclusions and Outlook

Ginsparg-Wilson fermions offer an attractive framework to tackle one of the most notorious and long-standing problems in the physics of hadrons. We have completed the first part of our project to investigate the mechanism of the $\Delta I = 1/2$ rule and found evidence for a significant non-perturbative enhancement of the amplitude A_0 in the artificial situation of a light and degenerate charm quark. This represents indirect evidence for the scenario that a substantial part of the enhancement should be due to the decoupling of the charm quark. An important cross-check of these findings, based on correlators saturated with zero modes is under way. As the latter method is numerically cheaper, it will be advantageous to study the case of a heavy charm quark also in this framework. The additional correlation functions have already been programmed, and also the required chiral corrections are being worked out[19].

References

1. L. Giusti, P. Hernández, M. Laine, P. Weisz and H. Wittig, *A strategy to study the rôle of the charm quark in explaining the* $\Delta I = 1/2$ *rule*, JHEP **0411**, 016, 2004.
2. P. Hernández and M. Laine, *Charm mass dependence of the weak Hamiltonian in chiral perturbation theory*, JHEP **0409**, 018, 2004.
3. L. Giusti, P. Hernández, M. Laine, C. Pena, J. Wennekers and H. Wittig, $K \to \pi\pi$ *Amplitudes from Lattice QCD with a Light Charm Quark*, Phys. Rev. Lett. **98**, 082003, 2007.
4. P. Hernández and M. Laine, *Probing the chiral weak Hamiltonian at finite volumes*, JHEP **0610**, 069, 2006.
5. P. Dimopoulos, L. Giusti, P. Hernández, F. Palombi, C. Pena, A. Vladikas, J. Wennekers, H. Wittig, *Non-perturbative renormalisation of left-left four-fermion operators with Neuberger fermions*, Phys. Lett. B **641**, 118, 2006.
6. P. Hernández, M. Laine, C. Pena, E. Torró, J. Wennekers and H. Wittig, *Weak low-energy couplings from topological zero-mode wavefunctions*, PoS **LAT2007**, 356, 2007, arXiv:0710.5647 [hep-lat].
7. M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *Light Quarks and the Origin of the* $\Delta I = 1/2$ *Rule in the Nonleptonic Decays of Strange Particles*, Nucl. Phys. B **120**, 316, 1977.
8. J. Gasser and H. Leutwyler, *Thermodynamics of chiral symmetry*, Phys. Lett. B **188**, 477, 1987; *Spontaneously broken symmetries: Effective Lagrangians at finite volume*, Nucl. Phys. B **307**, 763, 1988.
9. H. Leutwyler and A. Smilga, *Spectrum of Dirac operator and role of winding number in QCD*, Phys. Rev. D **46**, 5607, 1992.
10. P. Hernández and M. Laine, *Correlators of left charges and weak operators in finite volume chiral perturbation theory*, JHEP **0301**, 063, 2003.
11. P. Hernández, K. Jansen, L. Lellouch and H. Wittig, *Non-perturbative renormalization of the quark condensate in Ginsparg-Wilson regularizations*, JHEP **0107**, 018, 2001.
12. M. Lüscher, *Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation*, Phys. Lett. B **428**, 342, 1998.
13. P.H. Ginsparg and K.G. Wilson, *A remnant of chiral symmetry on the lattice*, Phys. Rev. D **25**, 2649, 1982.
14. H. Neuberger, *Exactly massless quarks on the lattice*, Phys. Lett. B **417**, 141, 1998; *More about exactly massless quarks on the lattice*, ibid. **427**, 353, 1998.
15. L. Giusti, Ch. Hoelbling, M. Lüscher and H. Wittig, *Numerical techniques for Lattice QCD in the* ϵ *regime*, Comput. Phys. Commun. **153**, 31, 2003.
16. T. DeGrand and S. Schaefer, *Improving meson two-point functions in lattice QCD*, Comput. Phys. Commun. **159**, 185, 2004.
17. L. Giusti, P. Hernández, M. Laine, P. Weisz and H. Wittig, *Low-energy couplings of QCD from current correlators near the chiral limit*, JHEP **0404**, 013, 2004.
18. J. Wennekers and H. Wittig, *On the renormalized scalar density in quenched QCD*, JHEP **0509**, 059, 2005.
19. Th. Tilli, Diploma Thesis, University of Mainz, 2008 (expected); B. Knippschild, Diploma Thesis, University of Mainz, 2008 (expected).

Materials Science

R. O. Jones

Institut für Festkörperforschung, Forschungszentrum Jülich
52425 Jülich, Germany
E-mail: r.jones@fz-juelich.de

It has almost become a cliché to observe that materials science is one of the great beneficiaries of changes in the landscape of scientific computing. The increase of computing power of several orders of magnitude over the past 20 years has been accompanied by improved methods and algorithms that allow us to calculate structural, cohesive, and magnetic properties of systems of many atoms and several elements without using any input from experiment. Furthermore, it is possible to perform calculations simultaneously on families of materials in the time required for a single system only a decade ago. In amorphous and other disordered systems, to mention just one example, the motion of individual atoms can be calculated in a detail that is often unattainable by experimental means. Of course, radioactivity and poisonous materials present no special problems for computer simulations. The density functional (DF) formalism is the basis of most calculations in condensed matter physics that have no adjustable parameters, and its early development led to the award of the Nobel Prize in Chemistry in 1998. The DF theory is the basis of the five materials science contributions in this volume.

I am writing these lines on the day that Albert Fert and Peter Grünberg were awarded the 2007 Nobel Prize for Physics for their discovery of giant magnetic resonance (GMR). The change in resistivity of multilayer structures caused by an external magnetic field is the most spectacular (and technologically the most successful) example of *spintronics*. In their contribution, Bergqvist and Dederichs discuss possible future applications in this area based on dilute magnetic semiconductors, such as $Ga_{1-x}Mn_xAs$. Their focus is on the critical behaviour, and they combine the DF method to determine the exchange interactions in a model system that is then studied with Monte Carlo statistical methods. Magnetic properties are also discussed in the paper of Gruner, Rollmann, and Entel, who studied clusters of hundreds of transition element atoms. Their results for Fe_{561} and for Fe-Pt and Co-Pt alloys indicate that calculations for 1000-2000 atom systems are certainly feasible in the near future.

Magnetic storage in computers is so familiar that we may think that there are no alternatives. This is not the case. All users of personal computers must have wondered about (and annoyed by) the delay between switching on a PC and it becoming operational. The reason is that normal random access memory (RAM) is "volatile", i.e. the content vanishes when power is switched off, and the operating system must be "booted" from the hard disk. Non-volatile memory is familiar to us in the form of USB sticks and memories used in digital cameras, but the limitations of such "flash memory" in scaling to smaller dimensions have stimulated the search for highly stable, high density alternatives. The materials of choice at the moment are tellurium-based alloys, and some of the most interesting are discussed in the article by Jaakko Akola and me.

For as long as I can remember, electronic structure calculations have promised to deliver detailed understanding of heterogeneous catalysis, without doubt an area of huge technological importance. The contribution of Huber *et al.* shows that we have made much progress along this path. Their focus is on the structure of palladium clusters (and their oxides) located on magnesium oxide surfaces. This work provides a further example of the ways in which detailed calculations aid the interpretation of experimental data. The final contribution to this section (by Ramos and Bechstedt) provides an example of how "standard" density functional calculations, which provide us with geometrical structures and cohesive properties, can be extended to properties of excited states, such as optical properties. This can be done by the *GW approximation* to the self-energy of the systems, which in this case are nanocrystallites of germanium and silicon with several hundred atoms and an array of defects.

The applications described in this section are just examples of materials science projects that are carried out on the supercomputers of NIC (JUMP and JUBL). Calculations on large clusters for simulation times that were unthinkable a few years ago are becoming almost routine, and the coming generations of computers conjure up prospects that are even more exciting. The techniques developed by scientists to study properties of the "materials" world have reached, for example, systems whose sizes are relevant for reactions involving biological molecules. Perhaps the biologists should watch out for us!

Dilute Magnetic Semiconductors

Lars Bergqvist and Peter H. Dederichs

Institute of Solid State Physics
Research Centre Jülich, 52425 Jülich, Germany
E-mail: {*l.bergqvist, p.h.dederichs*}*@fz-juelich.de*

We present first principles calculations for prospective magnetic materials for future applications in spintronics. Spintronics combine the charge and spin degree of freedom of the electrons making it possible to create novel devices with increased functionality compared to semiconductor devices used today. The material studied is focused on dilute magnetic semiconductors (DMS) like $Ga_{1-x}Mn_xAs$ that play a key role in semiconductor spintronics. Due to their ferromagnetic properties they can be used in magnetic sensors and as spin injectors. The basic problem for applications are, however, the relatively low Curie temperatures of these systems. We therefore focus on understanding the magnetic properties and on a reliable calculation of Curie temperatures from first principles. We have developed a theoretical framework for calculating critical temperatures by combining first principles calculations and statistical methods like Monte Carlo simulations.

1 Introduction

Magnetic materials have been known to mankind for a very long time. Perhaps the most famous example is the compass needle but today magnets are used everywhere, as permanent magnets in electrical motors in our cars, in harddrives on our computers and mp3-players, in loadspeakers, credit cards etc. Meanwhile, semiconductor based devices that use the charge of the individual electrons as the information carrier, have extensively been used in the last century in integrated devices like microprocessors, memory modules and diodes used in computers, amplifiers, cell phones etc. The performance of these devices has increased by several orders of magnitude over the last decades (according to Moore's law, the performance on transistors double every 18 mounts). This performance gain is due to, among other things, development in manufacturing and fabrication which makes it possible to shrink the dimension of the transistors and pack them more densely in order to make a faster device which simultaneously runs cooler and more efficiently. However, nowadays this route to improve performance seems to face severe difficulties, because the dimensions are becoming so tiny that classical physics no longer is applicable and quantum effects start to become important, like tunneling through the gate oxide layer causing malfunction of the transistor. Of course, this has been known for a long time and so far the industry has all the time found new clever ways to overcome the problem and make the devices go faster. For instance, Intel replaced the gate oxide from silicon dioxide to a rare earth based material (hafnium silicide) in their latest 45nm process, which improves not only the switching speed of the transistor but also reduces the power needed. However, all this development *will* eventually come to an end where it is simply not possible to shrink the dimensions further. It is here *spintronics* enter. The basic idea is to make devices where quantum mechanics is actually employed and not fought against. The electrons not only carry electrical charge but also spin, a purely (relativistic) quantum effect that can not be explained by classical physics. The spin is then manipulated in the spintronic device,

instead of the charge, schematically shown in Figure 1.

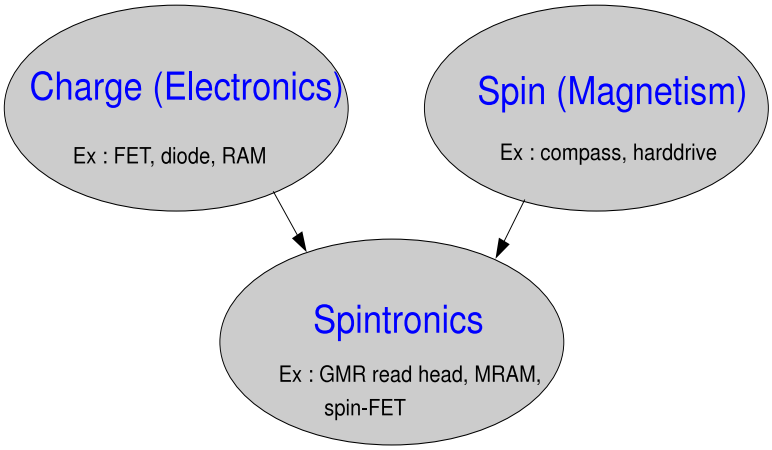


Figure 1. Schematic overview of spintronics which combines both charge (electronics) and spin (magnetism) into a novel field of research and applications.

The most successful application of spintronics to date is without any doubt the giant magneto resistance (GMR) read head in harddrives, an effect that was discovered by P. Grünberg and A. Fert who received the 2007 Nobel Prize in Physics. In very simple terms, the effect arises in a multilayer that consists of alternative magnetic and nonmagnetic materials in which the electrical resistivity depends on the magnetic configuration of the magnetic layers.

Dilute magnetic semiconductors like $Ga_{1-x}Mn_xAs$ are considered as ideal materials for spintronics. Only a few percent of Mn impurities are needed to achieve ferromagnetism. Thus the hope is that they can substitute the metallic ferromagnets to achieve all-semiconductor spintronics. Moreover they are half-metallic exhibiting a 100% spin polarization at the Fermi level E_F, being ideal for spin dependent devices. Their biggest disadvantage is that the reported Curie temperatures are well below room temperature. Here we will present some results of *ab initio* calculations, showing the complexity of these systems.

2 Computational Methods

The main goal of this study is to calculate critical temperatures T_c of several DMS systems. For applications the T_c needs to exceed room temperature by some margin since it determines in a way the upper working temperature of the device. To calculate T_c for realistic systems using only quantum mechanics methods is a formidable task which has not yet been solved. We have therefore used an alternative approach where we have split the problem into two parts, here called a two-step approach. In the first step, we apply quantum mechanical calculations using density functional theory to calculate the total energy of the system. These calculations do not have any adjustable parameters nor need

input from experiments. The total energy is then mapped to a simplified solvable model, in this case a classical Heisenberg model, which we solve in the second step by applying statistical methods to estimate T_c. Although simplified, this is by no means simple and in order to obtain reliable results, large scale calculations are called for.

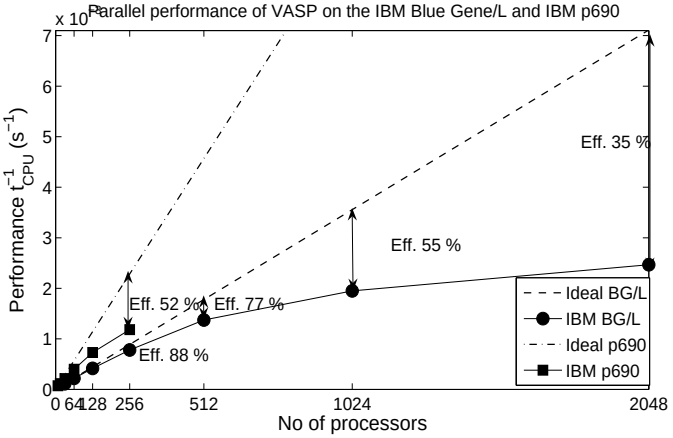


Figure 2. Parallel performance of VASP on the IBM p690 cluster (JUMP) and the IBM Blue Gene/L (JUBL) at the Research Centre Jülich.

The quantum mechanical calculations are based on density functional theory (DFT) in combination with the local spin density approximation (LSDA). These calculations are able to predict many physical quantities of real materials using only the lattice geometry and atomic numbers as input parameters. In practice, many different implementations exist based on various degrees of sophistication. We use the program VASP (Vienna *Ab initio* Simulation Package) of J. Hafner and G. Kresse[1–3] for our studies. VASP is a complete package for performing *ab initio* molecular dynamics (MD) simulations using either pseudopotentials or the projector-augmented wave method (PAW) and a plane wave basis set. The VASP code is spin polarized and fully relativistic and can therefore treat most systems in the periodic table. The program is written in Fortran90 and is massively parallelized using the MPI library. The parallelization is done over the bands and/or over plane wave coefficients to improve scaling. Moreover, the program relies a lot on Fast Fourier Transforms (FFT) which is also done in parallel. From a code optimizing point of view, the code steps through time, and performs a Fast Fourier Transform (FFT) and a matrix diagonalization using iterative routines (RMM-DIIS or blocked Davidson) every time step. The diagonalization is optionally performed using SCALAPACK routines. The VASP program has been ported to both the IBM p690 cluster (JUMP) and to the IBM Blue Gene/L (JUBL) at the Research Centre Jülich.

In Figure 2 we show a scaling plot of the parallel performance of VASP on the JUMP and the JUBL computers. The total number of atoms in the unit cell was 250, which is a typical number for a large-sized problem. The program on JUMP does scale acceptable up to 256 processors with the efficiency above 0.5. Due to the faster interconnect on the Blue Gene/L computer, the program scales much better than on JUMP and show acceptable

scaling of more than 50% up to 1024 processors. Given that each individual processor on JUBL is slower than on JUMP, one needs approximately twice as many processors on JUBL compared to JUMP for the same performance but this is a minor issue due to the larger resources and the relatively good scaling on JUBL. A profiling of the parallel performance on the JUBL reveals that the matrix diagonalizing routines which use the parallel SCALAPACK library do not scale well above 512 processors which hinders a good overall scaling (above 512 procs.). However, it should be said that the most time consuming routine in VASP, which is the iterative matrix diagonalizing routines using the RMM-DIIS algorithm, scales perfectly linearly up to 2048 processors. The memory requirement of VASP is by all means quite large, which has the implication that one usually needs to run the program in coprocessor mode on JUBL. However, the memory requirement per node decreases with increasing number of processors due to data distribution over the nodes. Moreover, with the new Blue Gene/P computer this should be a minor issue due to the larger available memory on that computer.

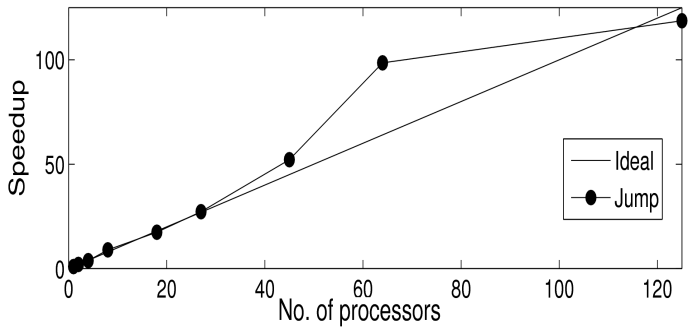


Figure 3. Speedup of parallel MC on the IBM p690 cluster (JUMP) at the Research Centre Jülich.

Once the classical Heisenberg model has been constructed from our first principles calculations, we apply statistical methods to obtain finite temperature properties from the calculated exchange coupling constants. We are using Monte Carlo (MC) simulations for this. MC gives a numerical solution in which both the positional disorder and the spin fluctuations are treated exactly. The simulations are performed on a fixed lattice using the Metropolis or the heat bath algorithm. The fixed lattice allows for fine grained parallelization using domain decomposition of the lattice using MPI which in turn allows for studies on very large simulation cells. Technical details are found in Ref.[4]. Magnetic exchange interactions as far out as 25 shells of neighbours from a central atom can be treated, which is crucial for real metallic systems where the interactions typically are long ranged as well as for diluted systems.

In Figure 3, scaling of the MC program as a function of number of processors on the JUMP computer are displayed. As seen from the figure, excellent scaling, in fact super-

linear scaling for some sizes, are obtained up to 125 processors. The program has a very small memory requirement (a few megabytes) and also run very well on the Blue Gene/L.

3 Diluted Magnetic Semiconductors (DMS)

A diluted magnetic semiconductor (DMS) can be realized by alloying of magnetic elements in a regular semiconductor like GaAs. Some of the Ga cations will then be replaced by magnetic atoms like Mn which carry a local magnetic moment. Moreover, if the valence of the substituted atom is different from the cation, holes are introduced which can mediate (ferro) magnetic interactions between the Mn atoms. An important consequence of this substitution is that the magnetic atoms are diluted and have a random distribution. DMS like $Ga_{1-x}Mn_xAs$ are one of the most hopeful materials for an all semiconductor spintronics. An enormous amount of papers and strong progress has been produced in this field, but the bottleneck for applications is still the availability of DMS with Curie temperature T_c above room temperatures. Thus understanding, predicting and realizing DMS with T_c larger than room temperature is one of the most important problems in spintronics. The highest conclusively reported T_c of DMS is around 170 K for 8% Mn doped GaAs, which is too low for applications.

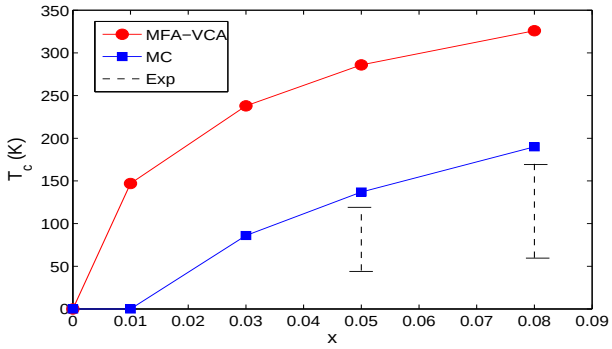


Figure 4. Calculated Curie temperatures of Mn-doped GaAs as a function of Mn concentration x using the mean field approximation (MFA-VCA) and Monte Carlo simulations (MC). The exchange coupling parameters are calculated from LDA.

Over the years, more and more knowledge has been gathered on DMS systems. For example, it has been demonstrated that electron doping strongly reduces the ferromagnetism leading to a disordered local moment state (DLM)[5]. The dominating exchange mechanism leading to ferromagnetic order, i.e. Zeners's $p-d$-exchange and Zener's double exchange, have been clarified by *ab initio* calculations and are now well understood[6,7]. Moreover, it has been demonstrated that percolation and disorder effects play a crucial role in these system by strongly reducing the ferromagnetism[8–10]. As an example, in Figure 4 calculated Curie temperature of Mn-doped GaAs is displayed as a function of concentration of substituted Mn-atoms employing the LDA. The most accurate method, namely MC simulations, yields T_c values that due to percolation problems are considerably lower than

previous mean field estimates (MFA-VCA), e.g. by Dietl *et. al.*, so that percolation seems to present a general obstacle for high T_c in DMS. This is particularly true for DMS like (Ga,Mn)N or (Zn,Cr)Te, in which due to the wide band gap the interaction is very short ranged. However, even for (Ga,Mn)As, where the interaction is fairly long ranged, this is a significant effect (see Figure 6).

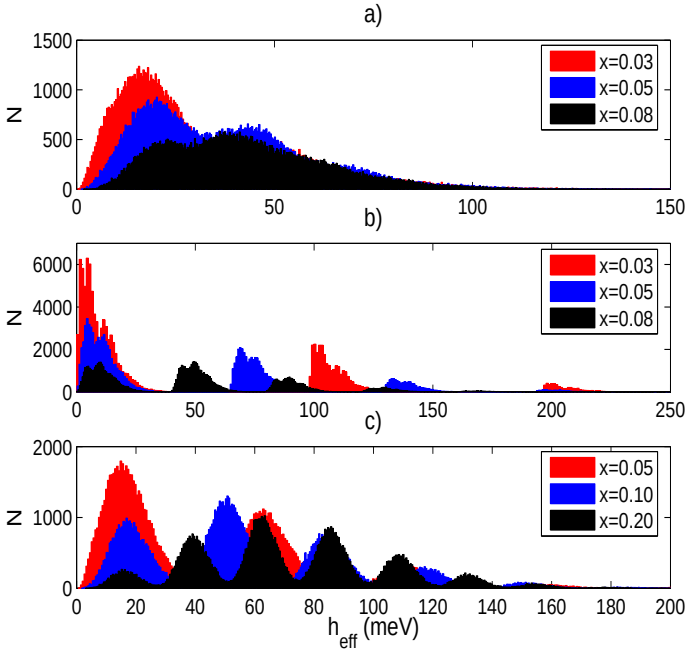


Figure 5. Distributions of local magnetic fields for a) Mn-doped GaAs, b) Mn-doped GaN and c) Cr-doped ZnTe. x denotes the concentration of magnetic impurities.

In order to get a more complete picture why the mean field approximation together with the virtual approximation (MFA-VCA) sometimes completely fails, but at other times gives reasonable results we have performed an analysis of the local effective magnetic fields h^{eff} in DMS materials. To be more specific, the h^{eff} is defined as

$$h_i^{\text{eff}} = \sum_j J_{0j} \langle S_j^z \rangle, \tag{1}$$

for each magnetic site i and at $T = 0K$, $\langle S_j^z \rangle = 1$. In a non-random system, each site i has the same value of h^{eff}. However, in contrast, in random system like DMS each site has a different local environment and we will instead obtain a distribution of local effective magnetic fields as displayed in Figure 5 for Mn-doped GaAs, Mn-doped GaN and Cr-doped ZnTe. The distribution is obtained for huge systems with around 10^5 magnetic impurities to make sure that basically all local environments are included. If a deeper analysis is performed, one can conclude that the MFA-VCA method is reasonable if the distribution is close to a Gaussian. If not, then the MFA-VCA method will fail spectacularly.

The distribution for Mn-doped GaAs (Figure 5a) does not differ dramatically from a Gaussian, especially for the higher concentrations. Consequently, of the three different DMS systems considered here it is the system where the MFA-VCA estimate is closest to the exact MC value (approximately a factor of 2 too large). The situation in GaN is very different (Figure 5b). The distribution is very far from a Gaussian but instead it consists of several peaks (arising from the dominant interactions in the 110-direction). It is clear from this distribution that the MFA-VCA estimate will be very wrong and this is indeed the case (the exact MC values are about one order of magnitude lower). An intermediate case is Cr-doped ZnTe, in which the distribution shows several smaller peaks. On the other hand, the concentration of magnetic atoms is larger than in the two cases above making the MFA-VCA estimate slightly improved.

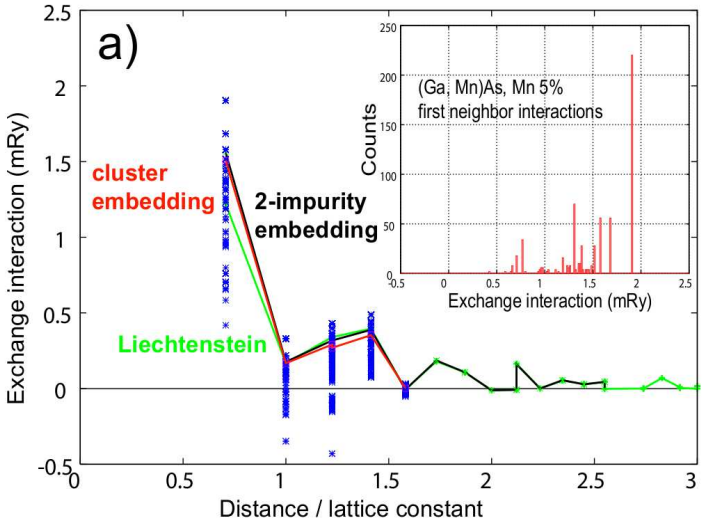


Figure 6. Exchange interactions in 5% Mn-doped GaAs using cluster embedding in CPA medium as a function of distance. The inset shows the distribution of nearest neighbour interaction.

Another important question is: How well can the exchange coupling constants J_{ij} between two impurities at sites i and j be described by a unique value J_{ij} independent of the local environments, i.e. the location of other magnetic impurities in the vicinity of the (i, j) pair? In the standard approach by Lichtenstein this effect is thought to be included by using averaged Green's functions based on the coherent potential approximation (CPA). We have investigated his problem in detail by calculating J_{ij} for a series of disordered local environments (clusters) embedded in the CPA medium. The results presented in Figure 6 show large fluctuations of the J_{ij} data, depending strongly of the positions of third and fourth Mn impurities in the neighbourhood of i and j. The inset shows the frequency distribution of nearest neighbour values J_{01}. The configurational average is in disagreement with the Lichtenstein's results based on the CPA, at least for the nearest

neighbour interactions. However, it agrees well with the exact results for two impurities i and j embedded in the CPA medium, showing that the disorder is well described by the CPA, provided that the 2-impurity problem is solved correctly.

4 Concluding Remarks

We have presented a theoretical investigation of prospective materials for future applications in spintronics, namely diluted magnetic semiconductors, using a combination of quantum mechanical calculations and statistical methods. Unfortunately, the critical temperatures are still too low for practical use.

Acknowledgments

L.B. acknowledges support from the EU within the framework of Marie Curie Actions for Human Resources and Mobility.

References

1. G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
2. G. Kresse and J. Furthmüller, Comput. Mat. Sci. **6**, 15-50 (1996).
3. `http://cms.mpi.univie.ac.at/vasp/`.
4. L. Bergqvist, Ph.D. Thesis, Department of Physics, Uppsala University (2005).
5. P. A. Korzhavyi, I. A. Abrikosov, E. A. Smirnova, L. Bergqvist, P. Mohn, R. Mathieu, P. Svedlindh, J. Sadowski, E. I. Isaev, Yu. Kh. Vekilov and O. Eriksson, Phys. Rev. Lett. **88**, 187202 (2002).
6. K. Sato, P. H. Dederichs and H. Katayama-Yoshida, Europhys. Lett. **65**, 403 (2003).
7. K. Sato, P. H. Dederichs, H. Katayama-Yoshida and J. Kudrnovský, J. Phys.:Cond. Matter **16**, S5491 (2004).
8. L. Bergqvist, O. Eriksson, J. Kudrnovský, V. Drchal, P. A. Korzhavyi and I. Turek, Phys. Rev. Lett. **93**, 137202 (2004).
9. L. Bergqvist, O. Eriksson, J. Kudrnovský, V. Drchal, A. Bergman and I. Turek, Phys. Rev. B **72**, 195210 (2005).
10. K. Sato, W. Schweika, P. H. Dederichs and H. Katayama-Yoshida, Phys. Rev. B **70**, 201202(R) (2004).

Large-Scale First-Principles Calculations of Magnetic Nanoparticles

Markus Ernst Gruner, Georg Rollmann, and Peter Entel

Fachbereich Physik, Universität Duisburg-Essen, 47048 Duisburg, Germany
E-mail: {Markus.Gruner, Peter.Entel}@uni-due.de

Modern supercomputers as the IBM Blue Gene/L provide the possibility to investigate large systems containing several hundred transition metal atoms. We present results on two examples, the size dependent evolution of structure and magnetism of elemental iron nanoparticles and the identification of structural comparison of competing morphologies of near-stoichiometric Fe-Pt and Co-Pt nanoparticles, which are currently discussed as media for future ultra-high density recording applications.

1 Introduction

Transition metal nanoparticles are of growing interest, from the fundamental science point of view[1,2] as well as for technological applications.[3,4] Atomistic simulations of materials properties of nanometer sized objects can easily be carried out by classical molecular dynamics simulations using empirical model potentials which already permit simulations on mesoscopic length scales. In many cases, however, the relevant properties are related to the electronic structure, so that a quantum-mechanical approach is necessary. The recent evolution of supercomputing power allows for an *ab initio* treatment of systems in the nanometer size regime. This comes at hand for technological applications where the miniaturization of functional units plays an important role. Within this contribution, we present two examples in which large scale *ab initio* calculations contribute to resolve open questions in this field.

2 Computational Method

The density functional theory (DFT) provides a suitable way to calculate the properties of systems containing several hundreds or even a few thousands of atoms from first principles, i. e. from quantum-mechanical grounds. According to the theorem by Hohenberg and Kohn,[5] the ground state of the electronic system can be uniquely described as a functional of the electron density. The resulting equations are solved in a self-consistency approach. The exact form of the exchange-correlation part of the functional is not known; nevertheless, this approach has proven its validity with overwhelming success.[6] During the last decades, the DFT approach has been implemented and consecutively refined in a vast number of commercial and freely available codes. In our calculations, we use the Vienna *Ab initio* Simulation Package (VASP),[7] which has proven in many examples to provide an excellent compromise between speed and accuracy, using the projector augmented wave approach[8] to describe the interaction with the nuclei and the core electrons. The clusters were placed in a cubic supercell, surrounded by a sufficient amount of vacuum to prevent interactions with the periodic images; k-space integration was restricted to the Γ-point. For

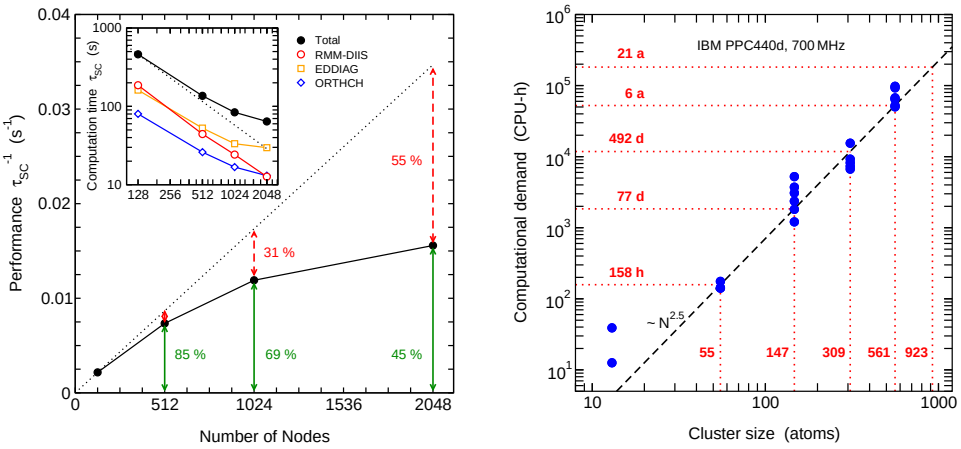


Figure 1. Left: Scaling behaviour of a nano-cluster with 561 Fe atoms on the IBM Blue Gene/L. The performance given by the inverse average computation time for an electronic self-consistency step, τ_{SC}^{-1}, is shown as a function of the number of nodes (black circles). The dashed lines describe the ideal scaling behaviour. The inset shows a double logarithmic plot of τ_{SC}. The open symbols in the inset refer to the scaling of the most time consuming subroutines. Right: Double logarithmic plot of the computation time per CPU needed on JUBL for a full geometric optimization of Fe-Pt nanoclusters of magic cluster sizes according to Eq. (1).

the exchange-correlation functional, we employed the generalized gradient approximation (for further details, refer to Refs. 9–12).

DFT calculations of large systems are not only time-consuming but also very demanding with respect to I/O bandwidth and memory. Thus the distribution over hundreds or thousands of CPUs is necessary to successfully complete such problems, especially on machines with hard limitations concerning CPU speed and memory per node as on the Blue Gene/L. Here, however, the threefold high-bandwidth, low-latency network helps to increase scalability so that large systems can be handled efficiently. Fig. 1 (right) shows as an example the scaling behaviour for an Fe_{561} cluster (4488 valence electrons) where on 1024 nodes still 70 % of the ideal (linear scaling) performance is reached. The largest calculations carried out by our group contained 8000 valence electrons and ran with sufficient efficiency on four Blue Gene/L racks (4096 nodes). The DFT approach does not only yield total energies and electron densities, also forces can be calculated accurately. This allows for quasi-static structure optimizations which we perform on the Born-Oppenheimer surface, i. e. in the electronic ground state of the system. This requires to fully converge the electronic system before making an ionic movement. In VASP, an efficient prediction scheme for the new wavefunctions is implemented. Nevertheless, a considerable number of electronic self-consistency steps – for larger systems typically several thousands – have to be performed to find one meta-stable minimum on the potential energy surface. Figure 1 (right) shows the evolution of the required computation time for a complete geometric relaxation as a function of the size of the system. A pragmatic fit to a power law yields an increase of the computation time with system size with an exponent of about 2.5.

The systematic scan of the potential energy surface in order to find the most stable morphologies is a very demanding task even for small clusters containing a few tens of

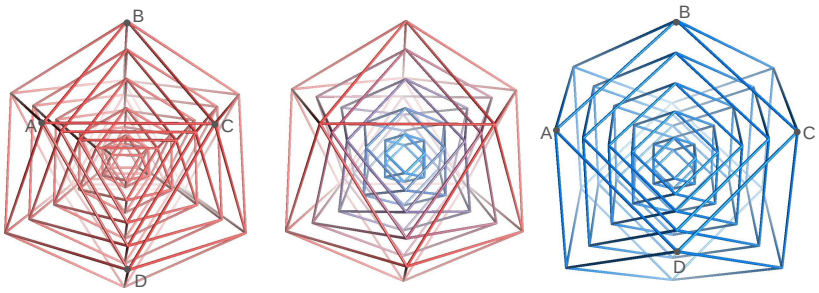


Figure 2. Edge models of the optimized Fe_{561} icosahedron (left), cuboctahedron (right) and SMT isomer (center). Separately for each shell, the corner atoms of the faces are connected by lines, the atoms themselves are omitted for clarity. For icosahedra and cuboctahedra, shells of the same shape but different sizes are stacked into each other. For the SMT isomer, the shape of the outermost shell is icosahedral and changes towards the inside continuously along the Mackay-path to a cuboctahedron, which represents the shape of the innermost shell. The Mackay transformation works by stretching the bond $\overline{AC}$ of the icosahedron and turning the two adjacent triangular faces ABC and ACD into the same plane producing the square face $ABCD$ of the cuboctahedron.

atoms. However, it becomes hopeless in the nanometer range. Therefore, a permissible way out is to restrict to an "educated guess" of selected morphologies. So-called magic-number clusters are a good starting point. Their size N is given as a function of the number n of closed geometric shells:

$$N = 1/3\left(10n^3 + 15n^2 + 11n + 3\right) = 13, 55, 147, 309, 561, 923, \ldots \qquad (1)$$

Magic cluster sizes appear to be particularly stable for the late $3d$ elements.[13] In addition, they allow for a comparison of several paradigmatic geometries: Cuboctahedra with a face centred cubic (fcc) structure, body centred cubic (bcc) isomers (Bain-transformed cuboctahedra), Mackay-icosahedra[14] and Ino-decahedra[15].

3 Structure and Magnetism of Elemental Iron Nano-Clusters

As a bulk material, iron belongs certainly to the most important elements in everyday life. Many application are related to its magnetic properties. Iron possesses a complex structural phase diagram, with bcc α- and δ-phases below 1185 K and above 1167 K, the fcc γ-phase in between and the hexagonal close packed (hcp) ϵ-phase at high pressures. Also, iron nanoparticles provide an interesting research field, too, due to their applicability in biomedicine (see, e. g., Ref. 16), but also from a fundamental point of view, since the evolution of their structure as a function of particle size has not been resolved so far: Transmission electron micrographs (TEM) of 6 nm particles suggest a bcc structure,[17] while for 13 atoms, first principles calculations predict a Jahn-Teller distorted icosahedron.[18] Recently the authors addressed this problems within a large scale *ab initio* approach which revealed that already above a crossover size of about 150 atoms the bulk-like bcc structure is energetically favored.[11] Furthermore, another important previously unreported structural motif was identified. The so-called shellwise Mackay transformed (SMT) isomer evolves by geometric relaxations from ideal icosahedra and cuboctahedra, if small distortions along the Mackay path[14] are imposed (cf. Fig. 2). Thus, for the case of iron, both geometric forms

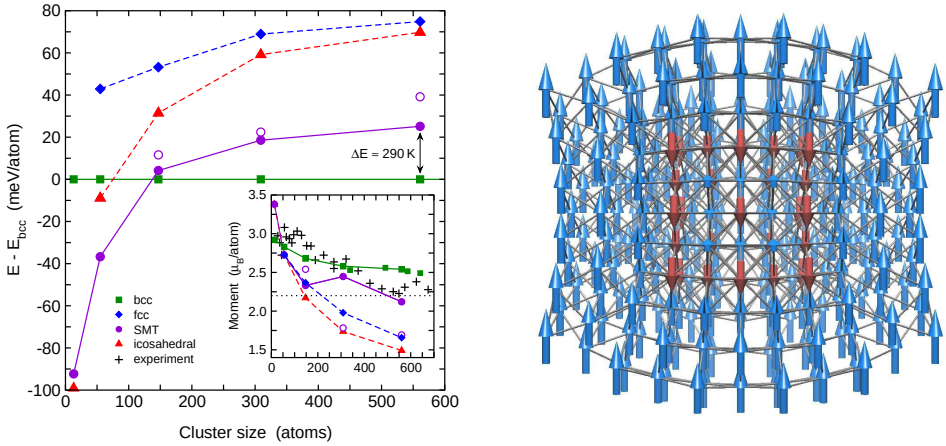


Figure 3. Left: Comparison of the energies of various isomers of elemental iron nano-clusters as a function of their size (data originally published in Ref. 11). The bcc isomer has been chosen as reference. Open symbols denote SMT isomers with magnetic configurations higher in energy. The lines are intended as guide to the eye, broken lines refer to instable structures. The inset shows the variation of the average magnetic moment of the different isomers with size (experimental moments taken from Ref. 1). Right: Spin configuration of the lowest energy Fe_{561} SMT isomer (cross section). The arrows refer in length and orientation to the atomic moments.

must be considered unstable even for small cluster sizes. The comparison of the total energies of the different morphologies in Fig. 3 shows that the SMT isomer is a hot candidate for the ground state of Fe_{55}. Even up to 561 atoms, the energy difference between the bcc and the SMT isomer remains – in contrast to the cuboctahedra and icosahedra – in the range of thermal energies; thus the occurrence of the SMT structure appears realistic at finite temperatures. The calculated magnetic moments can be directly compared with the measurements on small transition metal nanoparticles by Billas *et al.*[1] (inset of Fig. 3). In accordance with a previous DFT study on clusters of up to $N = 400$ atoms,[2] the moments of the bcc clusters agree very well with experiment in this size range. For larger clusters, however, considerable deviations occur. In fact, for $N = 561$, the SMT isomer shows the best agreement with the experimental magnetization data, owing to a shell-wise antiferromagnetic configuration of the cluster core.

The search for the origin of this unusual transformation requires an inspection at the atomistic level, which may become a tedious or even impossible task for large system sizes. In our case, help comes from the common neighbour analysis (CNA),[19] which has become a widely used tool for the identification of structural patterns in classical molecular dynamics studies of large systems. The CNA characterizes the local environment of an atom by a set of signatures, which can be compared with the characteristic result for an ideal bulk or surface structure. If one restricts to the first neighbour shell, signatures are obtained for each pair of neighbours, containing information on the number of nearest neighbours both atoms have in common, the number of bonds between these neighbours and the longest chain connecting them. Figure 4 shows the result of a CNA applied on optimized isomers. While the bcc and fcc clusters are uniform in structure, we find for the icosahedron the typical mixture of fcc, hcp and – along the five-fold symmetry axes –

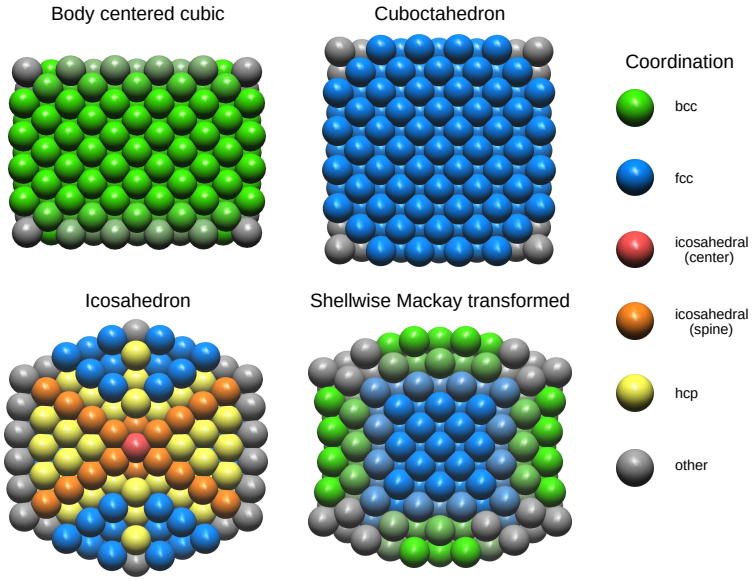


Figure 4. Cross sections of optimized geometries of 561 atom iron nanoparticles. The colour coding describes the local coordination of the atoms obtained by a common neighbour analysis. Bright colours refer to perfect matching of the CNA-signatures with the ideal bulk and surface configurations. With increasing number of deviations the colours turn into grey.

icosahedral environments. The SMT isomer, however, shows a rather unexpected pattern: While the central atoms are fcc coordinated according to the nearly completed Mackay-transformation of these shells, we find no trace of the typical icosahedral signatures in the outermost shells which, however, retain their overall icosahedral shape. Instead, their signatures show typical signs of an energetically favorable bcc-like coordination, accompanied by a pair distribution function which is rather typical for amorphous materials.[12]

4 Morphologies of Fe-Pt and Co-Pt Nanoparticles

In the field of ultra-high density magnetic recording, the long-lasting exponential increase in storage density over time still seems unbroken. While eight years ago, densities of 35 GBit/in^2 were state of the art, recent lab demos reached values around 400 GBit/in^2 and the expectations of the manufacturers go up to 10 to 50 TBit/in^2 in the future. The main obstacle to further miniaturization is the so-called superparamagnetic limit, which threatens the long-time stability of the information stored.[20] The Néel relaxation time of a recording media grain is decribed by an exponential dependence on the product of anisotropy constant and volume divided by temperature. This imposes a lower boundary for the possible size above which the magnetization of a grain is (for a sufficiently long time) not affected by thermal relaxation processes. One widely discussed strategy to overcome this problem is the use of regular patterned arrays of nanoparticles, where one bit is essentially represented by one particle.[3,4] To obtain sufficiently small particles with thermally stable

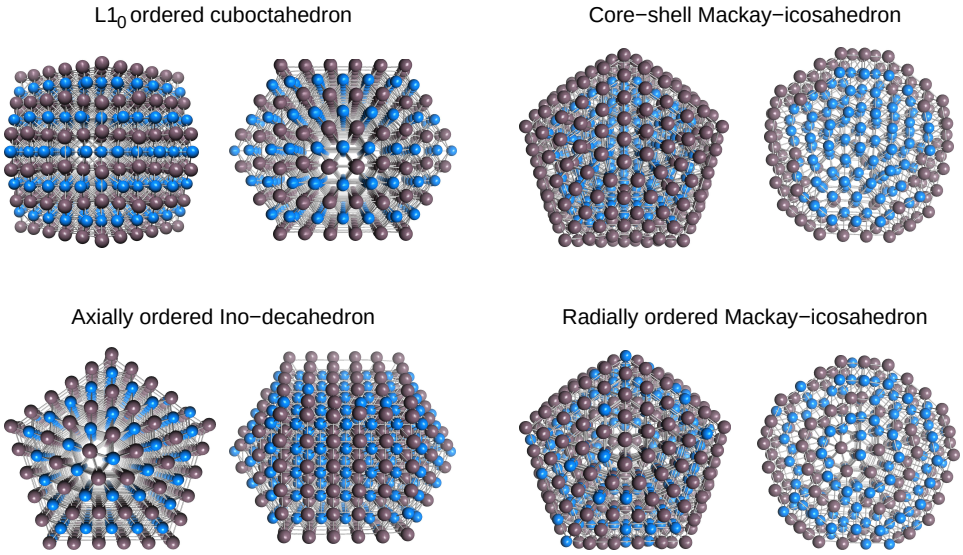


Figure 5. Various ordered morphologies of $Fe_{265}Pt_{296}$ (or $Co_{265}Pt_{296}$, respectively) nanoparticles after structural relaxation in two different aspects. The blue spheres denote Fe (or Co) atoms, the brownish spheres Pt. The icosahedra are shown in cross section to visualize their inner structure.

magnetizations, near-stoichiometric FePt and CoPt alloys in the slightly tetragonally distorted $L1_0$ phase are currently considered as the most promising materials, since their bulk anisotropy constants are about one order of magnitude larger than that of currently used media.[20] Considering these values, particle sizes of 4 nm or even smaller appear feasible.

The current production route is to fabricate disordered fcc FePt particles from gas phase experiments or wet-chemical production routes and obtain the ordered $L1_0$ phase in a further annealing step.[4] However, it was reported recently that $L1_0$ particles in the interesting size range with a sufficient magnetocrystalline anisotropy may be difficult to obtain in this way.[21–23] High resolution transmission electron microscopy (HRTEM) showed the occurrence of multiply twinned morphologies such as icosahedra and decahedra at diameters around 6 nm and below (e. g., 24, 25). These consist of several strained twins (20 in the case of Mackay-icosahedra and 5 for Ino-decahedra). Therefore such morphologies cannot be expected to exhibit a large uniaxial magnetocrystalline anisotropy even if the individual twins are perfectly $L1_0$-ordered, since they have different crystallographic orientations.

In order to shed more light on the origin of these problems, we started *ab initio* calculations of FePt and CoPt nanoparticles with of up 561 atoms (≈ 2.5 nm in diameter), comparing nearly 60 configurations of particles of both materials with different sizes.[26] Although the largest particles are still too small to avoid superparamagnetism, general trends can be formulated from the results, as the largest clusters already possess a balanced surface-to-volume ratio (45 % at 2.5 nm as compared to 32 % at 4 nm). We find that disordered phases are considerably higher in energy than the desired $L1_0$ phase. However, ordered multiply twinned morphologies are substantially more favorable throughout the investigated size range. The calculated energy differences can be as large as $\Delta E = E - E_{L1_0} = -30$ meV/atom in the case of the radially ordered $Fe_{265}Pt_{296}$ icosahe-

dron, which is still close to thermal energies, but sizeable -90 meV/atom for $Co_{265}Pt_{296}$. These results explain at least in part the experimental difficulties to obtain particles with large magnetocrystalline anisotropy at small sizes. The fully segregated icosahedron is a special case: It is with $\Delta E = +20$ meV/atom higher in energy than the $L1_0$ cuboctahedron in the case of $Fe_{265}Pt_{296}$ but with $\Delta E = -92$ meV/atom much more favorable in Co-Pt, pointing out strong segregation tendencies in the latter case.

5 Outlook

In this contribution, we wanted to demonstrate that large scale DFT calculations are a powerful tool providing information on material related properties on an atomistic level, which may not be obtained from experiment so far and can thus contribute significantly to the solution of unresolved questions in materials science and technology. Our ongoing work in the field of nanoparticles for ultra-high density recording applications includes the investigation of the size-dependence of the magnetocrystalline anisotropy energy of FePt and CoPt nanoparticles and the exploration of possible strategies to stabilize the $L1_0$ phase by co-alloying with a third element. Considering the possibilities provided by the new Blue Gene/P at Jülich Supercomputing Centre (JSC), *ab initio* investigations of nanoparticles of 3.5 or even 4 nm, containing 1415 and 2057 atoms according to Eq. (1) and being of relevant size for technological applications, appear feasible in near future.

Acknowledgments

We thank the John von Neumann Institute for Computing for granting computation time and support and, especially, P. Vezolle of IBM and I. Gutheil of JSC for their substantial efforts, making the VASP code run efficiently on the Blue Gene/L system. Financial support was granted by the Deutsche Forschungsgemeinschaft through SPP 1239 and SFB 445.

References

1. I. M. L. Billas, A. Châtelain, and W. A. de Heer, *Magnetism from the atom to the bulk in iron, cobalt, and nickel clusters*, Science, **265**, 1682, 1994.
2. M. L. Tiago, Y. Zhou, M. M. G. Alemany, Y. Saad, and J. R. Chelikowsky, *Evolution of magnetism in iron from the atom to the bulk*, Phys. Rev. Lett., **97**, 147201, 2006.
3. S. Sun, C. B. Murray, D. Weller, L. Folks, and A. Moser, *Monodisperse* FePt *nanoparticles and ferromagnetic* FePt *nanocrystal superlattices*, Science, **287**, 1989, 2000.
4. S. Sun, *Recent advances in chemical synthesis, self-assembly, and applications of* FePt *nanoparticles*, Adv. Mater., **18**, 393, 2006.
5. P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev., **136**, B864, 1964.
6. W. Kohn, "Electronic structure of matter – wave functions and density functionals", in: Nobel Lectures, Chemistry 1996-2000, I. Grenthe, (Ed.), p. 213. World Scientific, Singapore, 2003.
7. G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B, **54**, 11169, 1996.

8. G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*, Phys. Rev. B, **59**, 1758, 1999.
9. M. E. Gruner, G. Rollmann, S. Sahoo, and P. Entel, *Magnetism of close packed* Fe_{147} *clusters*, Phase Transitions, **79**, 701, 2006.
10. M. E. Gruner, G. Rollmann, A. Hucht, and P. Entel, *Massively parallel density functional theory calculations of large transition metal clusters*, Lecture Series on Computer and Computational Sciences, **7**, 173, 2006.
11. G. Rollmann, M. E. Gruner, A. Hucht, P. Entel, M. L. Tiago, and J. R. Chelikowsky, *Shell-wise Mackay transformation in iron nano-clusters*, Phys. Rev. Lett., **99**, 083402, 2007.
12. M. E. Gruner, G. Rollmann, A. Hucht, and P. Entel, *Structural and magnetic properties of transition metal nanoparticles from first principles*, vol. 47 of *Advances in Solid State Physics*, Springer, Berlin, in press.
13. M. Pellarin, B. Baguenard, J. L. Vialle, J. Lerme, M. Broyer, J. Miller, and A. Perez, *Evidence for icosahedral atomic shell structure in nickel and cobalt clusters - comparison with iron clusters*, Chem. Phys. Lett., **217**, 349, 1994.
14. A. L. Mackay, *A dense non-crystallographic packing of equal spheres*, Acta Cryst., **15**, 916, 1962.
15. S. Ino, *Stability of multiply-twinned particles*, J. Phys. Soc. Jpn., **27**, 941, 1969.
16. D. L. Huber, *Synthesis, properties, and applications of iron nanoparticles*, Small, **1**, 482, 2005.
17. T. Vystavel, G. Palasantzas, S. A. Koch, and J. Th. M. De Hosson, *Nanosized iron clusters investigated with in situ transmission electron microscopy*, Appl. Phys. Lett., **82**, 197, 2003.
18. P. Bobadova-Parvanova, K. A. Jackson, S. Srinivas, and M. Horoi, *Density-functional investigations of the spin ordering in* Fe_{13} *clusters*, Phys. Rev. B, **66**, 195402, 2002.
19. H. Jónsson and H. C. Andersen, *Icosahedral ordering in the Lennard-Jones liquid and glass*, Phys. Rev. Lett., **60**, 2295, 1988.
20. D. Weller and A. Moser, *Thermal effect limits in ultrahigh-density magnetic recording*, IEEE Trans. Magn., **35**, 4423, 1999.
21. B. Stahl, J. Ellrich, R. Theissmann, M. Ghafari, S. Bhattacharya, H. Hahn, N. S. Gajbhiye, D. Kramer, R. N. Viswanath, J. Weissmüller, and H. Gleiter, *Electronic properties of* 4-nm FePt *particles*, Phys. Rev. B, **67**, 014422, 2003.
22. T. Miyazaki, O. Kitakami, S. Okamoto, Y. Shimada, Z. Akase, Y. Murakami, D. Shindo, Y. K. Takahashi, and K. Hono, *Size effect on the ordering of* $L1_0$ *FePt nanoparticles*, Phys. Rev. B, **72**, 144419, 2005.
23. O. Dmitrieva, B. Rellinghaus, J. Kästner, M. O. Liedke, and J. Fassbender, *Ion beam induced destabilisation of icosahedral structures in gas phase prepared FePt nanoparticles*, J. Appl. Phys., **97**, 10N112, 2005.
24. Z. R. Dai, S. Sun, and Z. L. Wang, *Shapes, multiple twins and surface structures of monodisperse FePt magnetic nanocrystals*, Surf. Sci., **505**, 325, 2002.
25. D. Sudfeld, O. Dmitrieva, N. Friedenberger, G. Dumpich, M. Farle, C. Song, C. Kisielowski, M. E. Gruner, and P. Entel, *HR-TEM studies of* FePt *nanoparticles by exit wave reconstruction*, Mater. Res. Soc. Symp. Proc., **998E**, 0998–J01–06, 2007.
26. M. E. Gruner, G. Rollmann, P. Entel, and M. Farle, *Multiply twinned morphologies of* Fe-Pt *and* Co-Pt *nanoparticles*, submitted.

Structural Patterns in Ge/Sb/Te Phase-Change Materials

Jaakko Akola[1,2] **and R. O. Jones**[1]

[1] Institut für Festkörperforschung, Forschungszentrum Jülich, D-52425 Jülich
E-mail: {j.akola, r.jones}@fz-juelich.de

[2] Nanoscience Center, Department of Physics, FI-40014 University of Jyväskylä, Finland

Phase-change (PC) materials are widely used for optical recording and computer memory, but the structure of the phases involved and the nature of the phase transition in the nanoscale bits pose continuing challenges. Massively-parallel density functional simulations of the amorphous structure of the prototype materials $Ge_2Sb_2Te_5$ (GST) and GeTe show that there is long-ranged order among Te atoms in both. The crucial structural motif is a four-membered ring with alternating atoms of types A (Ge, Sb) and B (Te), an '$ABAB$ square'. The rapid amorphous-to-crystalline phase change can be viewed as a re-orientation of disordered $ABAB$ squares to form an ordered lattice, and vacancies (voids) in the disordered phases to provide the necessary space. The improved phase change performance of GST over GeTe can be correlated with the higher vacancy concentration (11.8% cf. 6.4%). Supercomputers with the power of the IBM Blue Gene/L are essential for density functional calculations of this scale and precision.

1 Introduction

Modern computers and other electronic devices place great demands on the density, speed, and stability of memory. Phase change (PC) materials already play important roles in rewritable media (CD-RW, DVD-RW, DVD-RAM, Blu-ray disc) and are prime candidates for wider applications in the future.[1] The basis of their function is the rapid and reversible transition between the crystalline and amorphous forms in nanoscale bits (~ 100 nm) of an alloy.[2] The latter arises from quenching after a localized and short (~ 1 ns) laser annealing to a temperature above the melting point T_m. Longer laser heating (~ 50 ns) to above the glass transition temperature but below T_m leads to a metastable crystal. Changes in the optical and electronic properties provide the means to monitor the transition.

The most common PC materials are Te-based alloys. The $Ge_xSb_yTe_{1-x-y}$ family provides prototypes, and $Ge_2Sb_2Te_5$ (GST) is already a commercial success in DVD-RAM (random access memory). Essential to our understanding of the properties of these materials is a knowledge of the structures of the different phases, but these are difficult to determine in binary or ternary alloys with large numbers of vacancies. Even the structure of the ordered phase of GST is controversial: Yamada[3] proposed that the metastable phase has a rock salt structure with 'Na' sites occupied randomly by Ge and Sb atoms and vacancies, and 'Cl' sites by Te. However, Kolobov et al.[4] proposed that Ge and Sb atoms are displaced from their ideal positions, enabling the order-disorder transition to occur as an 'umbrella flip' of Ge atoms from octahedral to tetrahedral positions. Most recently, x-ray fluorescence holography of an epitaxial layer of GST indicated a cubic structure with tetrahedral site symmetry about Ge atoms.[5] It is astonishing that PC materials could become the basis of commercially successful products with so much uncertainty about the structures of the phases involved.

Recent extended x-ray absorption fine structure (EXAFS) measurements on GST have found significant concentrations of Ge-Ge bonds and indications that overcoordinated Te

atoms play a role,[6] and there has been renewed focus on the role of vacancies.[7] The reverse Monte Carlo (RMC) analysis of synchrotron radiation data from amorphous (a-) GST and a-GeTe suggests that the ring structure of the former is dominated by four- and six-membered rings also seen in the crystal.[8]

DF calculations are free of adjustable parameters, but their demands on computational resources have restricted simulations on GST systems to relatively small unit cells (less than 200 atoms in all cases, less than 100 in most) and time scales that are often much shorter than those relevant experimentally. The present calculations are much larger in both number of atoms (460 for GST, 216 for GeTe) and time scale (hundreds of picoseconds). Our simulations start from liquids at 3000 K and do not favor particular crystal types. Further details of all aspects, including the electronic structure, are provided in Ref. 9.

2 Numerical Methods

2.1 Density Functional Calculations

The DF calculations were performed with the Car-Parrinello molecular dynamics package (CPMD)[10] using Troullier-Martins[11] pseudopotentials and the generalized gradient approximation of Perdew, Burke and Ernzerhof (PBE)[12] for the exchange-correlation energy functional E_{xc}. We use periodic boundary conditions, usually with one point (**k**=0) in the Brillouin zone, and the kinetic energy cutoff of the plane wave basis set is 20 Ry.

The GST simulations adopted the densities of the metastable crystalline (6.35 g/cm^3) and amorphous (5.88 g/cm^3) phases and were based on the crystalline (rock salt) structure with 512 atomic sites, where the Na and Cl sites are occupied by Ge/Sb atoms (20 % each) and vacancies (10 %), and Te atoms (50 %), respectively. The Na sites are populated randomly with Ge, Sb, and vacancies, so that the sample contains 460 atoms (102 Ge, 102 Sb, 256 Te) and 52 vacancies. The cubic cells for the c-GST and a-GST have sides of 24.05 and 24.62 Å, respectively. The GeTe system contains 216 atoms at the density of the amorphous phase (5.61 g/cm^3), and the box size is 18.61 Å. Ge, Sb, and Te are relatively heavy atoms, and we use Born-Oppenheimer MD, for which the time steps adopted (6.050 and 3.025 fs for initialization and data collection, respectively) are much longer than in the Car-Parrinello approach.

Memory of the crystalline starting structure was erased by starting the simulations at 3000 K (liquid), and details of the cooling in GST to the melting point (900 K) and to 300 K and 100 K are given in Ref. 9. Cooling and data collection in both GST and GeTe took well over 300 ps. Such relaxation times are essential for simulating structural transitions and place great demands on computing power.

2.2 Data Analysis

MD methods allow us to follow the coordinates R_i and velocities v_i of all atoms, and insight into the local order can be found from the distributions of the bond (θ_{ijk}) and dihedral angles (γ_{ijkl}). The radial distribution function (RDF) or pair correlation function $g(r)$ is a spherically averaged distribution of interatomic vectors,

$$g(r) = \frac{1}{\rho^2} \left\langle \sum_i \sum_{i \neq j} \delta(r_i)\delta(r_j - r) \right\rangle , \quad (1)$$

where ρ is the density. Partial radial distribution functions $g_{\alpha\beta}(r)$ are calculated by restricting the analysis to the elements α and β. The local structure can also be characterized by average coordination numbers found by integrating $g_{\alpha\beta}(r)$ to the first minima $R_{\min}$

$$n_{\alpha\beta} = \int_0^{R_{\min}} dr\ 4\pi r^2 \rho_{\alpha\beta}(r) g_{\alpha\beta}(r). \tag{2}$$

We calculate the structure factor $S(Q)$ by Fourier transforming the $g_{\alpha\beta}(r)$ to give the partial structure factors $S_{\alpha\beta}(Q)$, which are weighted according to the atomic fractions c_α and form factors $f_\alpha(Q)$ (for x-rays x) or the Q-independent coherent scattering lengths b_α (for neutrons n). In discussing the topology of nearest neighbours, it is convenient to separate the atoms into types A (Ge, Sb) and B (Te). It is also essential to give precise definitions of the regions where there are few atoms (vacancies, voids, cavities). Full details and the angular distribution functions are provided in Ref. 9.

Dynamical information includes the velocity autocorrelation function C_v:

$$C_v(t) = \frac{1}{N} \sum_{i=1}^{N} \frac{\langle v_i(0) \cdot v_i(t) \rangle}{\langle v_i(0) \cdot v_i(0) \rangle}, \tag{3}$$

where N is the number of particles. The self-diffusion constant for all atoms or those of species α can be determined from C_v or from the coordinates $R_\alpha(t)$.[9]

3 Results

3.1 Radial Distribution Functions and Structure Factors

The partial radial distribution functions (RDF) are shown in Fig. 1. The Te-Te curves [Fig. 1(a)] show that a-GeTe and a-GST have long-range correlations of Te atoms up to 10 Å. There are few Te-Te bonds (0.1 and 0.3 per Te atom for a-GeTe and a-GST) and the second maximum at 4.16 Å dominates. The region below 7 Å shows parallels to the ordered structure (rock salt). There are differences at larger distances, and the cubic Te sublattice proposed in the 'umbrella flip' model[4] is absent. The homopolar RDF for the remaining atoms (Ge, Sb: type A) exhibit little structure beyond ~ 6 Å.

GeTe and GST form mainly Ge-Te (and Sb-Te) bonds with pronounced first maxima in the partial RDF [Fig. 1(b)]. The Ge-Te and Sb-Te bonds are shorter than in the crystal. The location of the first minimum usually marks the limit of bond lengths, but the minima in a-GeTe and a-GST move to $3.6 - 3.9$ Å, so that there are many Ge-Te (and Sb-Te) pairs with 'intermediate' separations. The average coordination number of Te is near 3 in both a-GeTe and a-GST.

Amorphous GST has Ge-Ge, Sb-Sb, and Ge-Sb bonds with coordination numbers (Eq. 2) 0.4, 0.6, and 0.2, respectively. The EXAFS value for Ge-Ge bonds is 0.6 ± 0.2.[6] Such bonds are more common in a-GeTe, where our coordination number (1.1) compares well with an EXAFS measurement (1.2).[13] All total coordination numbers are larger than given by the '$8 - N$ rule' (Ge: 4, Sb: 3, Te: 2), and the coordination of Te atoms is lower in a-GST. The total coordination numbers are 3.4 in a-GST and 3.7 in a-GeTe.

The structure factors $S(Q)$ of a-GST have been calculated at 300 K (Fig. 2) and 900 K. There are minor differences between the $S(Q)$ calculated for neutrons and x-rays, and the

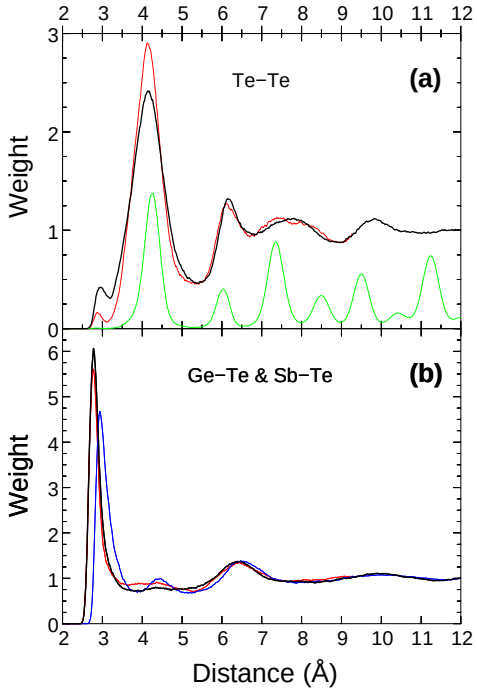


Figure 1. Radial distribution functions of a-GST and a-GeTe at 300K. (a) Te-Te. Black: a-GST, red: a-GeTe, green (scaled by 0.25): metastable GST crystal at 300 K. (b) Partial RDF. Black: Ge-Te bonds in a-GST, red: Ge-Te bonds in a-GeTe, blue: Sb-Te bonds in a-GST.

curves calculated from the total $g(r)$ (where the elements have equal weight) are also very similar. The calculated peak positions are at slightly lower Q-values than those measured, but all features are resolved. The structure factor for the amorphous phase has more structure and sharper peaks than in the liquid phase (see Ref. 9). There are more homopolar bonds (less AB alternation) and more disorder in the latter.

3.2 Ring Structures

Irreducible rings (the shortest closed loops) can be used to characterize the crystalline-amorphous phase transition, and the distributions in a-GST and a-GeTe (Fig. 3) differ: a-GST has a pronounced maximum for $n = 4$, the weight of larger rings decreases, and there is a pronounced odd-even alternation. 86% of the four-membered rings have $ABAB$ alternation, in both a-GST and a-GeTe. The bond angle distributions are peaked around 90°, and we denote these units as '$ABAB$ squares'. $60 - 80$% of Ge atoms participate in at least one such square depending on the bond cutoff distance chosen (3.2–3.4 Å).

The large number of alternating four-membered rings is evident in Fig. 4(a), where we highlight the relevant atoms and bonds in a-GST. We note the presence of $ABAB$ squares (and cubes) [Fig. 4(b)] and the relationship to the ordered $ABAB$ squares of the rock salt phase. Close to the cube in the centre of the cell there are intermediate distances where a slight re-orientation would increase the number of AB bonds and $ABAB$ squares.

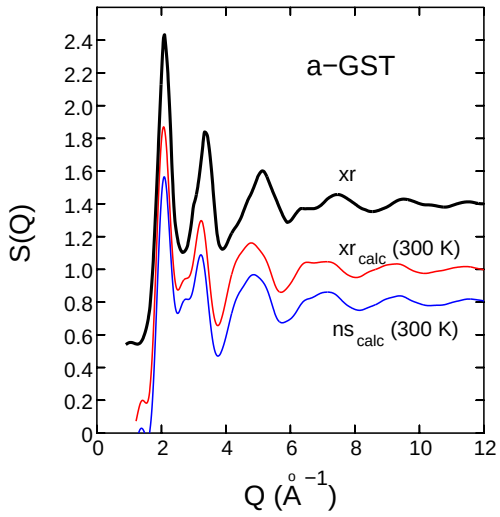


Figure 2. Structure factor $S(Q)$ of a-GST. The curve of Kohara et al. (xr, Ref. 8) is displaced by 0.4.

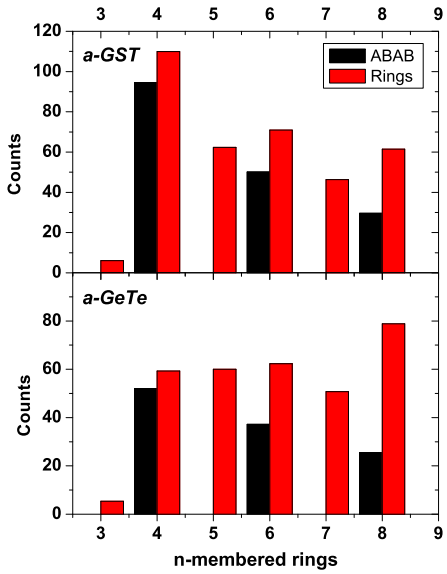


Figure 3. Statistics of irreducible n-fold ring configurations of a-GST and a-GeTe at 300 K. Black: corresponding alternating AB configurations (A: Ge/Sb, B: Te).

3.3 Vacancies

Vacancies play important roles in PC materials. The cavity analysis reveals a wide range of volumes and shapes, with a total volume of 11.8% and 6.4% for a-GST and a-GeTe,

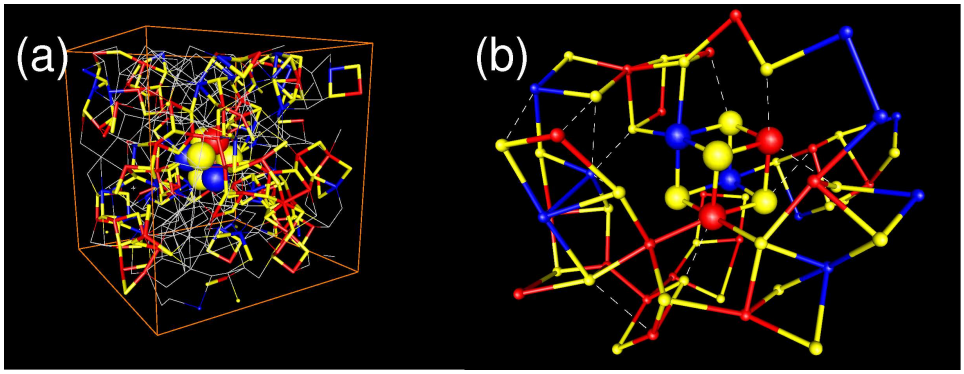


Figure 4. $ABAB$ squares and cubes in a-GST. (a) Simulation box of a-GST (24.6 Å, 460 atoms) with atoms and bonds of $ABAB$ squares and cube highlighted. Red: Ge, blue: Sb, yellow: Te. (b) Local environment of $ABAB$ cube. Dashed lines mark intermediate distances ($3.2 - 4.0$ Å).

respectively. There are then numerous vacancies in a-GeTe, whereas the ideal crystal has none. The vacancies are surrounded mainly by Te atoms, as in the crystalline phases of pseudobinary GeSbTe compounds, and a medium-sized cavity in a-GST is shown in Fig. 5(a). Sharp corners and edges are common.

The volume distribution of cavities in a-GST and the vacancy-vacancy correlation function (inset) are plotted in Fig. 5(b). There is a bias towards small cavities, and in larger vacancies (50-100 Å^3) di- and multivacancies are important and protrusions are common. There are long-range vacancy-vacancy correlations to at least 10 Å (Fig. 5(b), inset) that are similar to the behaviour in Te. Comparison of amorphous (300 K) and liquid GST (900 K) shows that the latter contains more small cavities (total volume 13.8%), but the distributions of large cavities are almost identical. Correlations are evident at 900 K (to $\sim$ 7 Å), and the dominant peak is shifted from 5.1 (300 K) to 5.7 Å (900 K). The rapid amorphous-to-crystalline transition can be viewed as a vacancy-supported reorientation of $ABAB$ squares. Numerous AB bonds are formed, but few homopolar bonds need to be broken in order to achieve crystalline order.

3.4 Diffusion

The mean square displacements of different elements at 900 K and the corresponding linear fits have been calculated. The diffusion coefficients D_α are 3.93×10^{-5} (Ge), 4.67×10^{-5} (Sb), and 3.78×10^{-5} cm^{2}s^{-1} (Te). The coordination number of Sb in the melt (3.4) is lower than in a-GST, and it is the most mobile element. The diffusion coefficients of GeTe at 1000 K are 4.65×10^{-5} (Ge) and 3.93×10^{-5} cm^{2}s^{-1} (Te), and the viscosity ranges between $1.1 - 1.2$ cP for both alloys, depending on the particle radius chosen in the Stokes-Einstein relation. The measured viscosity of liquid GeTe ranges from ~ 1.9 cP at 1000 K to 1.3 cP at 1150 K.[14]

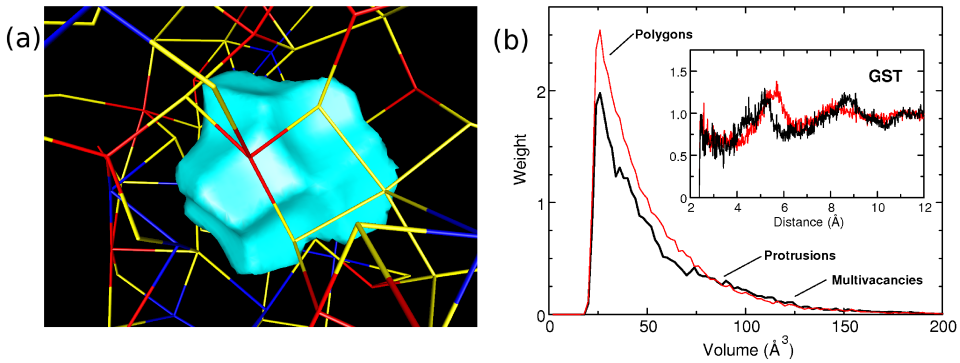


Figure 5. (a) A medium-sized cavity. (b) Volume distribution of cavities in a-GST, and (inset) vacancy-vacancy correlation function. Red: liquid at 900 K, black: a-GST at 300 K.

4 Discussion

Density functional calculations combined with molecular dynamics have been performed for liquid and amorphous GeTe and $Ge_2Sb_2Te_5$. The simulations involve 216 and 460 atoms, respectively, in the unit cell, over hundreds of picoseconds, i.e. a time scale similar to the quenching process from the liquid. Such calculations are impossible without computers with the power of the IBM Blue Gene/L. Amorphous GeTe and GST show long-range ordering of Te atoms and a high degree of alternating four-membered rings ('$ABAB$ squares') that are the main building blocks of both. Since the crystalline (rock salt) phase comprises perfectly ordered $ABAB$ squares, the rapid amorphous-to-crystalline transition can be seen as a re-orientation of $ABAB$ squares to form additional AB bonds and cubic subunits in a locally 'distorted octahedral' structure. Vacancies play a crucial role in providing the necessary space.

There are several reasons for the improved PC properties of GST relative to GeTe:

(1) *Vacancies.* The rhombohedral and fcc forms of GeTe are vacancy-free, and a-GeTe contains 6.4% vacancies by volume, about half of the value for a-GST. The local similarity between the amorphous and ordered phases is greater in GST than in GeTe due to the higher concentration of vacancies.

(2) *Homopolar bonds and AB ordering.* There are more Ge-Ge bonds in a-GeTe (1.1) than A-A bonds in a-GST (0.6 for Ge, 0.8 for Sb). This leads to different AB ordering and ring statistics, and $ABAB$ squares are more evident in a-GST.

(3) *Interatomic distances.* The first minima of the radial Ge-Te and Sb-Te distribution functions of a-GST extend to 3.8 to 3.9 Å, and these 'intermediate' distances make possible new AB bonds. This feature is less pronounced in a-GeTe, where there are fewer vacancies and the minimum is at 3.6 Å.

(4) *Atomic mobility.* The molten phases of a-GST (at 900 K) and a-GeTe (at 1000 K) have different atomic mobilities. This is due to the Sb atoms, whose diffusion constants are $\sim 20\%$ larger than those of Ge and Te.

The development of PC memory devices has favored chalcogenide semiconductors from the beginning, but current research and development focuses almost entirely on Te-

based alloys. We are now carrying out density functional simulations on both Te and the eutectic alloy $Ge_{15}Te_{85}$ in the amorphous and liquid states. The simulation cells contain up to 700 atoms, and the calculations require the power of the Blue Gene/L or its successor, Blue Gene/P.

Acknowledgments

J.A. acknowledges financial support from European Union project ULTRA-1D (NMP4-CT-2003-505457). The calculations were performed on IBM Blue Gene/L and p690 computers in the Forschungszentrum Jülich and made possible by grants from the FZ Jülich and the John von Neumann Institute for Computing (NIC).

References

1. See, for example, A. L. Greer and N. Mathur, Nature (London) **437**, 1246, 2005.
2. S. R. Ovshinsky, Phys. Rev. Lett. **21**, 1450, 1968.
3. N. Yamada, Mat. Res. Soc. Bull. **21**, 48, 1996.
4. A. Kolobov, P. Fons, A. I. Frenkel, A. I. Ankudinov, J. Tomonaga, and T. Uruga, Nat. Mater. **3**, 703, 2004, and references therein.
5. S. Hosokawa, T. Ozaki, K. Hayashi, N. Happo, M. Fujiwara, K. Horii, P. Fons, A. V. Kolobov, and J. Tominaga, Appl. Phys. Lett. **90**, 131913, 2007.
6. D. A. Baker, M. A. Paesler, G. Lucovsky, S. C. Agarwal, and P. C. Taylor, Phys. Rev. Lett. **96**, 255501, 2006.
7. See, for example, T. Matsunaga, R. Kojima, N. Yamada, K. Kifune, Y. Kubota, Y. Tabata, and M. Takata, Inorg. Chem. **45**, 2235, 2006.
8. S. Kohara, K. Kato, S. Kimura, H. Tanaka, T. Usuki, K. Suzuya, H. Tanaka, Y. Moritomo, T. Matsunaga, N. Yamada, Y. Tanaka, H. Suematsu, and M. Takata, Appl. Phys. Lett. **89**, 201910, 2006.
9. J. Akola and R. O. Jones, Phys. Rev. B **76**, 235201, 2007.
10. CPMD V3.11 Copyright IBM Corp 1990-2006, Copyright MPI für Festkörperforschung Stuttgart 1997-2001, (`http://www.cpmd.org`).
11. N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993, 1991.
12. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865, 1996.
13. Y. Maeda and M. Wakagi, Jpn. J. Appl. Phys. **30**, 101, 1991.
14. V. M. Glazov, S. N. Chizhevskaya, and N. N. Glagoleva, *Liquid Semiconductors* (Plenum, New York, 1969), p. 205.

Density Functional Modeling of Magnesia Supported Pd Clusters: On the Road to an Understanding of Nanoscale Catalysis

B. Huber[1,2], P. Koskinen[1,3,4], H. Häkkinen[4], and M. Moseler[1,2,3]

[1] Faculty of Physics, University of Freiburg, Herrmann-Herder-Str. 3, 79104 Freiburg

[2] Freiburg Materials Research Center, Stefan-Meier-Str. 21, 79104 Freiburg

[3] Fraunhofer Institute for Mechanics of Materials IWM, Wöhlerstr. 11, 79108 Freiburg
E-mail: mos@iwm.fhg.de

[4] Department of Physics, NanoScience Center, University of Jyväskylä

Understanding and predicting heterogeneous catalysis remains one of the main motivations underlying the science of supported nano-cluster. Surface science experiments provide important insights into nano-particles catalysed reactions. Often however, this progress is only achieved in combination with quantum-chemical atomistic simulations. Here we show how density functional theory can be used to understand experimental size evolutionary patterns in the activity of metal-oxide supported Pd clusters. We provide theoretical evidence that the reaction of supported Pd_N with molecular oxygen results in the formation of nano-oxides which are in epitaxy with the ceramic support. These oxides serve as a Mars-van-Krevelen oxygen reservoir and therefore play an important role in the catalyzed combustion of carbon monoxide.

1 Introduction

Oxide-supported transition metal clusters and nanoparticles have recently attracted significant attention due to their important role as components of model-catalysts[1–6], sensors[7], solar-cells[8] and magnetic recording devices[9]. For small clusters, functionality and structure are closely interrelated. However, knowledge of the structure of the bare cluster is insufficient since the interaction with the chemical environment might cause drastic structural changes. Here we show by *ab initio* simulations based on the density functional theory that the reaction with molecular oxygen transforms small, non-crystalline, magnesia-supported Pd-clusters to crystalline Pd_xO_y nano-oxide clusters that are in epitaxy with the underlying support. Restructuring of the Pd backbone is controlled by the electrostatic interaction with magnesia leading to a strong reduction of the O_2 dissociation barrier. The supported Pd_xO_y clusters are likely to serve as Mars-van-Krevelen[10] oxygen reservoirs in catalytic oxidation reactions as observed previously for PdO overlayers[11] and demonstrated here for the oxidation of CO molecules.

The pronounced chemical activity of small metal clusters is due to a combination of several factors, with their relative contributions strongly depending on cluster size and elemental composition[3,12]. For transition metal clusters, the highest occupied valence orbital is generally close to or within a manifold of d-derived states, and the average position and the width of such d-band dictates much of the characteristics of adsorption of molecules via covalent bonding[13]. In gas-phase palladium clusters, the closed-shell $4d^{10}5s^0$ atomic configuration opens via significant s-d hybridisation, which induces a spin-magnetic moment[14]

and enhances the reactivity of small clusters. When adsorbed on alkaline supports, such as magnesia, surface defects serve as strong trapping centres for the clusters that maintain their open valence shells and finite spin-moments[15]. Model-catalytic experiments and *ab initio* simulations indicated that even a single Pd atom adsorbed at a surface colour centre (FC) of the magnesia support can be catalytically active for CO oxidation with a direct CO_2 formation from either molecularly adsorbed oxygen or from a Pd-carbonate-complex[16].

This contribution reports an extensive density functional theory (DFT) study of the oxidation of magnesia-supported Pd_N clusters (N=4-9)[17,18] revealing a novel, unexpected low-temperature oxidation mechanism, involving a phase of an epitaxially grown Pd_xO_y/MgO(FC) that functions as a catalytic centre for CO oxidation.

2 The Born-Oppenheimer-Spin-Density-Molecular-Dynamics-Method

We employ the work horse of large scale quantum chemistry namely Kohn's density functional theory[19,20]. This method plays a dominant role for large systems since it reduces the complicated many electron system to a more tractable picture of a single electron in the mean field of the other electrons resulting in a three-dimensional eigenvalue equation, the so called Kohn-Sham equation.

$$\left(-\frac{1}{2}\nabla^2 + v_{eff,\sigma}(\mathbf{r}) \right) \phi_{i,\sigma}(\mathbf{r}) = \epsilon_{i,\sigma}\phi_{i,\sigma}(\mathbf{r}). \tag{1}$$

Here the $\phi_{i,\sigma}$ are a set of a single particle electronic wave function, $\epsilon_{i,\sigma}$ their energies and the effective potential is given by

$$v_{eff,\sigma}(\mathbf{r}) = v(\mathbf{r}) + \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{xc,\sigma}(\mathbf{r}). \tag{2}$$

The electron density n of the system as the central quantity of density functional theory derives from the occupied Kohn-Sham orbitals $n(\mathbf{r}) = \sum_{i,\sigma}^{occ} |\phi_{i,\sigma}(\mathbf{r})|^2$. In order to make the computations less expensive only chemically active electrons are considered and therefore a pseudo potential v is used for the confinement of the valence electrons representing the influence of the naked ions and the core electrons[21]. The so called exchange-correlation potential v_{xc} takes into account many body effects that are not included in the classical Coulomb field $\int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$ in the above equation. It is treated in the framework of the local density approximation (LDA) and improved by the generalized gradient approximation[22]. The spin of the system is explicitely taken into account by calculating the wave functions of both spin manifolds $\sigma = \uparrow, \downarrow$ and thus the description of magnetism is possible within this formalism. For more details on spin density functional theory, the reader is referred to standard text books[20].

For large systems, this equation can be solved with great accuracy and efficiency using a plane wave basis for the single electron wave functions[23]. However, the required memory and CPU speed still exceeds modern serial hardware and thus massive parallel computing is the only way to solve the Kohn-Sham equations for a large number of atoms. For instance, for the description of a Pd_{13} cluster on an MgO(001) surface[15], we utilized 3D grids of the order of 200 grid points in each dimension for the Fourier and real space representation of approximatelly 200 electronic wave functions resulting in a memory consumption of the

order of 10 Gbyte! Thus the demanding computation of nanocatalysts could only be done after reception of a generous computer grant for the JUMP at the NIC Jülich.

The method for the numerical solution of eq. (1) utilizes the Born-Oppenheimer-local-spin-density-molecular-dynamics (BO-LSD-MD) approach of Barnett and Landman[23] and benefits from the fact that the differential operator $-\frac{1}{2}\nabla^2$ is a simple multiplication by $-\frac{1}{2}k^2$ for the Fourier transform $\phi_{\mathbf{k}}$ of the wave function. An iterative Block-Davidson eigenvalue solver only needs the action of the Hamiltonean $-\frac{1}{2}\nabla^2 + v_{eff}$ onto a wave function and therefore a dual space technique treating the kinetic energy in Fourier and the potential energy part in real space provides a very efficient scheme to solve eq. (1). A domain decomposition of both spaces and an efficient parallelisation of the fast fourier transform (FFT) conecting k- and real space results in a very good parallel efficiency on massively parallel machines like the NIC JUMP. The FFT is also used to calculate the Coulomb field $\int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$ since it satisfies Poissons equation which is algebraic and thus easily solvable in k-space. After the solution of the Kohn-Sham equations, the forces on the ions are calculated employing the Hellmann-Feynman-Theorem[20]. For more details on the numerical aspects of the method see[23].

The atomic and electronic structure of the active part of the supported Pd_N/MgO system (comprising the vicinity of the colour centre of the MgO(100) surface, the adsorbed Pd_N cluster, and the reactant molecules O_2 and CO) were calculated within the DFT in combination with the BO-LSD-MD (for details concerning pseudopotential and plane wave cutoff see ref.[17]). The MgO surface with the F-centre is modelled by a two-layer *ab initio* cluster $Mg_{13}O_{12}$, embedded in a point-charge lattice to model the long-rage Madelung potential. 2087 alternating charges of +2 and -2 of Mg and O ions, respectively, were used. In addition, those positive point charges that would be nearest neighbours to the periphery O atoms of the $Mg_{13}O_{12}$ cluster were replaced by empty Mg pseudopotentials (i.e. those Mg ions are represented by the same pseudopotential as the ones inside the active region, but they do not contribute electrons to valence) in order to prevent unphysical polarization of O ions. The lattice parameter of the embedding part is fixed to the experimental lattice constant (4.21 Å) of bulk MgO. For each cluster system Pd_N/MgO(FC) as well as for the further calculations with the reactant molecules, the Mg and O ions of the substrate are kept fixed to the ideal truncated bulk positions. The palladium part and the molecules are treated fully dynamical during steepest-descent-like optimisations starting from many initial configurations. The cluster's spin degree of freedom S was determined by a series of total energy $E(S_z)$ calculations with restricted spin z component (S_z). In general, $E(S_z)$ showed a plateau with a slight energy variation of 100 meV for $S_z = S$ followed by a sudden rise of several hundred meV for larger S_z (a similar behaviour was observed previously for gas-phase Pd clusters, see Ref.[14]). Consequently, we estimate that the accuracy of the total energies and hence the reported adsorption energies regarding the spin degree of freedom is within 0.1 eV. The minimum energy pathways and barriers of the O_2 dissociation reaction were calculated employing the nudged elastic band method[24] by using typically 5-9 images along the reaction path. The CO reaction pathways were obtained by constrained energy minimisation for a set of fixed distances between the carbon and the closest oxygen on the nano-oxide. Atomic charges and local magnetic moments were determined from the accumulated electron spin density within the domains of a Richards-Voronoi decomposition of space[25] using the ionic radii of the atoms.

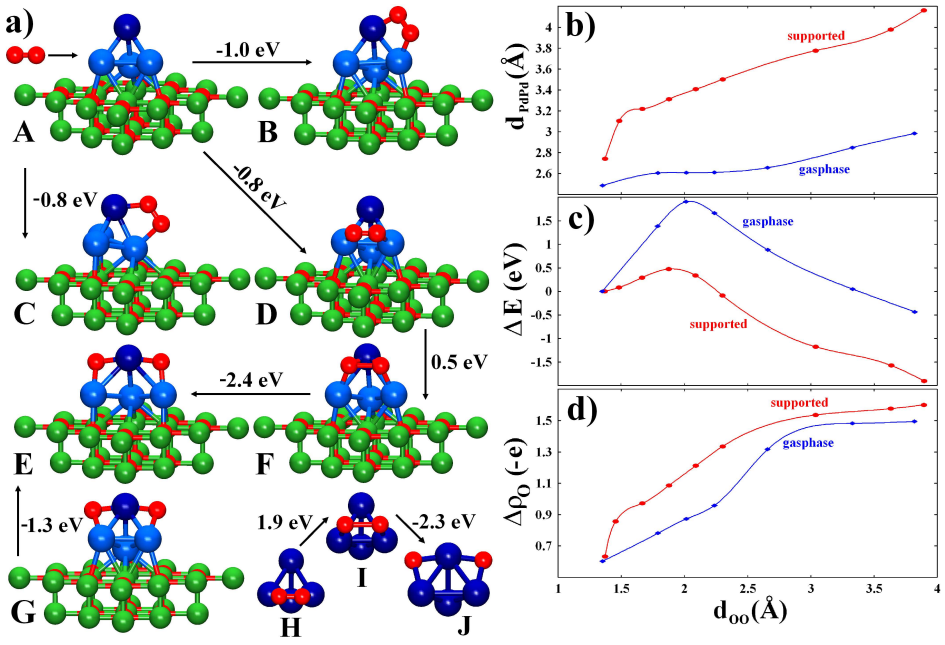


Figure 1. The oxidation of magnesia-supported Pd_4. Panel a) A: structure of the bare supported Pd_4; B-D: Pd_4 with moleculary adsorped O_2; E: ground state of Pd_4O_2/MgO(FC) with dissociated O_2; F: transition state between D and E. Oxygen atoms are represented by red, Mg by green, Pd atoms in contact with the substrate by light-blue and Pd in the second adlayer by dark-blue spheres. Arrows and numbers indicate transitions and energy differences between states. Panels b)-d): Nudged elastic band calculation of the pathways C→E and H→J. Shown are the Pd-Pd distance d_{PdPd}, the energy difference ΔE relative to the initial states and the charge accumulation $\Delta\rho_0$ on the oxygen as a function of the O-O distance d_{O-O}.

3 Results

The structures of the MgO-supported clusters were taken from an earlier DFT study[15] of Pd_N soft-landing on oxygen vacancies. The FCs represent the most prominent defects on MgO surfaces[27] and serve as pinning centres for the closest Pd atom (see Figs. 1a and 2). The other Pd atoms prefer neighbouring Mg-O bridges or form a second layer (light and dark blue atoms in Fig. 2, respectively).

The interaction of Pd_4/MgO(FC) with O_2 results in a barrierless metastable molecular adsorption on one of the three inequivalent Pd-Pd bridges with adsorption energies (E_{ad}) ranging between 0.8-1.0 eV (B-D in Fig. 1a). In the ground state (GS) of O_2Pd_4/MgO(FC), O_2 dissociates in the 2nd adlayer (E in Fig. 1a) with E_{ad} of 2.7 eV and a low dissociation barrier of 0.5 eV (Fig. 1a, F) with respect to a metastable molecular state (Fig. 1a, D). Note, that this activation barrier is considerably lower than the corresponding 1.0 eV barrier for room temperature O_2 dissociation on Pd(111)[28] implying that the reaction on Pd_N should be experimentally realised below room temperature. Interestingly, Pd atoms on the Mg-O bridges were displaced towards surface oxygen top positions creating a registry of Pd and O with the underlying support (E in Fig. 1a). This strong relaxation (as reflected in Pd-Pd distance in Fig. 1b) accompanies the dissociation indicating the importance of the

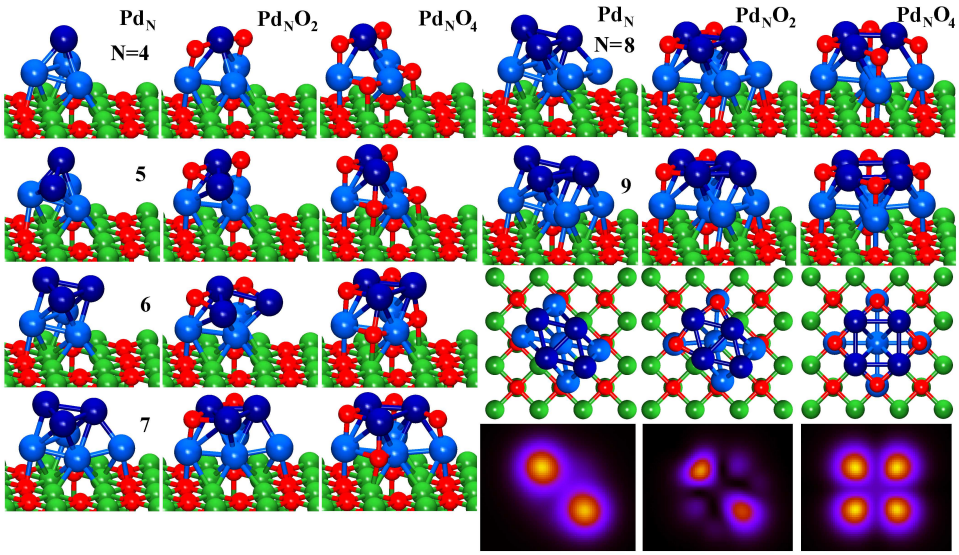


Figure 2. Structural growth motif of bare and oxidised Palladium clusters on MgO(FC). Panel in the right bottom corner: Tersoff-Hamann[26] STM simulation of supported O_xPd_9 assuming constant scan height of 2.5 Å above the cluster and contribution of orbitals within 0.2 eV below the Fermi energy.

structural fluxionality of the cluster. In fact, if Pd_4 were treated static during O_2 dissociation, the activation barrier would be 1.2 eV (see Fig.1a G for the final structure of the static Pd_4). Notably, no significant relaxation was observed for the oxidation of gas phase Pd_4[29] (Fig. 1a, reaction H→I→J). The cluster retained its tetrahedral shape during the reaction with a rather high activation energy of 1.9 eV (see Fig. 1b and 1c for a comparison of Pd-Pd distances and barriers pertaining to the gas phase and the supported reaction). This observation suggests that intra-cluster chemistry is unlikely to be the main driving force for the observed fluxionality.

The separation of the two oxygen atoms is accompanied by a significant electron transfer from Pd to O increasing the anionic charge from -0.6 e/O-atom for the molecular O_2 to -1.6 e/O-atom for the dissociated O_2 on the supported cluster and to -1.5 e/O-atom on the gas phase Pd_4 (Fig. 1d). Consequently, the supported oxidised Pd_4O_2 is influenced by a significant Madelung potential from the ionic substrate. We estimated the difference of the Madelung energies of the unrelaxed (Fig. 1a, G) and the relaxed Pd_4O_2 (Fig. 1a, E) employing a classical point charge model with charges and atomic positions taken from our DFT calculations. Within this model, the cluster gained 2.1 eV upon relaxation. This result compares rather well with the DFT energy difference of 1.3 eV (considering the neglect of bond breaking in the classical model).

Our calculations of the dissociation of M=1,2 O_2 molecules on the larger Pd_N clusters (N=5-9) revealed essentially the same processes as already discussed for Pd_4 resulting in a systematic growth pattern of $O_{2M}Pd_N$/MgO(FC) (Fig. 2). For all cluster sizes, the first pair of oxygen atoms was inserted in the 2nd adlayer epitaxially on O-sites of the magnesia with E_{ad} ranging between 2.7 eV and 3.4 eV. Notably, the optimum positions of the second

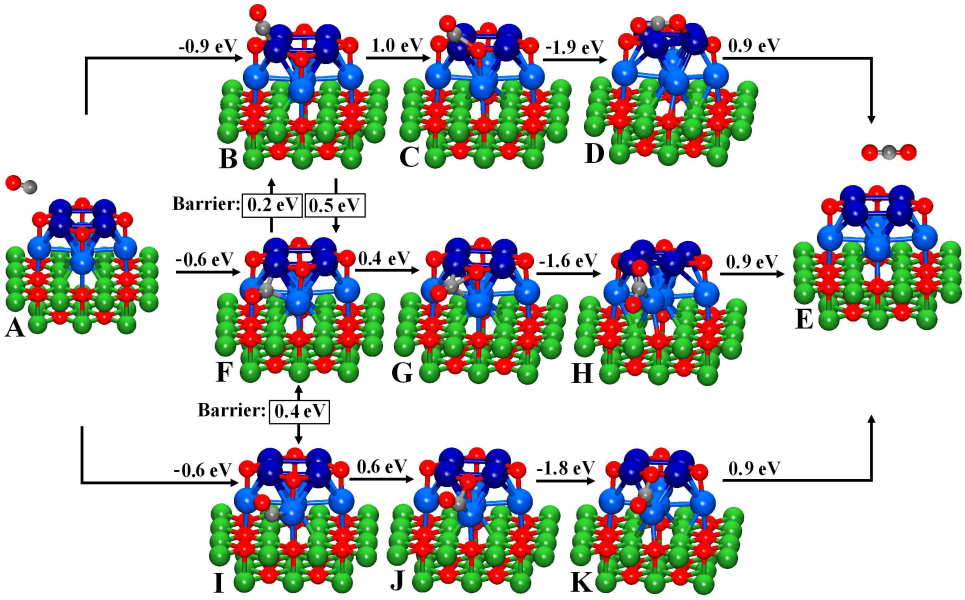


Figure 3. Possible CO oxidation reaction pathways A-B-C-D-E, A-F-G-H-E and A-I-J-K-E on O_4Pd_9/MgO(FC). Carbon is depicted as grey spheres. Structures (except transition states C, G and J) represent stable local minima. The transition state energies between B and F and between F and I are reported in boxes. After adsorption on B, the CO can take the direct route B-C-D (barrier 1.0 eV) or the indirect routes B-F-G-H or B-F-I-J-K having a multiple-barrier structure (0.5 eV/0.4 eV for the former and 0.5 eV/0.4 eV/0.6 eV for the latter) leading to effective barriers of 0.7-0.9 eV.

O pair showed a pronounced size-dependence. For smaller clusters (N=4-6) the 3rd and 4th oxygen atom occupy ideal MgO lattice sites in the first adlayer with reduced E_{ad} ranging between 1.6 eV and 1.8 eV. A structural transition occurs after Pd_6 leading to an O_4Pd_7/MgO(FC) with three oxygen atoms in the 2nd adlayer (E_{ad}=2.2 eV) and a complete depopulation of the 1st adlayer for Pd_8 and Pd_9 (E_{ad}=3.0 and 3.2 eV). The formation of a highly symmetric C_{4v} GS for O_4Pd_9/MgO(FC) produced a 2nd adlayer that is in perfect epitaxy with the MgO surface (see top view in Fig. 2). In principle, the transition from the C_{2v} Pd_9/MgO(FC) to the C_{4v} O_4Pd_9/MgO(FC) could be detected experimentally by scanning tunnelling microscopy (STM). In our simulated STM images (bottom row in Fig. 2) the rearrangement of the Pd atoms is clearly visible. Surprisingly, the square-like Pd_4O_4 unit in the completed 2nd adlayer represents the basic building block of a Pd(111) surface oxide[30]. Note however, that Pd(100) supports the formation of a $\sqrt{5} \times \sqrt{5}R27^\circ$ structure which has been interpreted as a (101) monolayer of tetragonal bulk PdO[31].

Recently, it has been demonstrated experimentally that PdO overlayers on Pd(100) can serve as Mars-van-Krevelen oxygen reservoirs for high temperature CO oxidation[11]. To elucidate the existence of a corresponding mechanism on Pd nano-oxides, we performed extensive DFT calculations of CO adsorption and CO_2 formation on O_4Pd_9/MgO(FC). The CO molecules adsorb preferentially on top of 2nd layer Pd atoms (structure B in Fig. 3 with E_{ads}=0.9 eV), on the Pd-Pd bridges (F in Fig. 3 with E_{ads}=0.6 eV) and at the side of the cluster close to the surface (I in Fig. 3 with E_{ads}=0.6 eV). The adsorption sites B,

F and I are interconnected by rather low barriers ($\leq$ 0.5 eV) indicating that CO molecules should be quite mobile on O_4Pd_9/MgO(FC) below room temperature.

We first describe CO_2 formation assuming low CO coverage i.e. the molecules get oxidized one at a time. At finite temperature, CO can be found at any of the three identified adsorption sites B, F or I. Site B can be reached either directly from gas-phase or through the metastable sites F and I that first trap CO via reverse spillover[32]. The reaction barrier for forming CO_2 via the direct route B-C-D is clearly the highest. The calculated barrier of 1.0 eV roughly corresponds to the lowest barriers for CO oxidation on bulk Pd(111)[33]. However, effective barriers of the order of only 0.7 to 0.9 eV exist for the indirect reactions B-F-G-H or B-F-I-J-K where both pathways have a multiple-barrier structure (Fig. 3). While the detailed comparison of reaction routes B-C-D, B-F-G-H and B-F-I-J-K would require quantitative information of the respective kinetic prefactors, it is reasonable to conclude that at low temperatures the indirect mechanisms B-F-G-H and B-F-I-J-K are likely to dominate. We note that the preformed CO_2 molecules are still relatively strongly bound to the catalyst (0.9 eV for each of the structures D, H or K), however, from the energetic point of view the reaction heat ($\geq$ 1.6 eV) should be sufficient to detach them to the gas phase[16].

At higher CO partial pressures, all the four equivalent CO adsorption sites B are populated. This full occupation of the first Pd-O layer makes the bridge adsorption site (F) unstable for CO, which in effect closes the reaction channel F-G-H and decouples channels B-C-D and I-J-K, but does not strongly modify the reaction barriers along these channels. Under this condition, the reaction can still quite effectively proceed along the low-barrier channel I-J-K for those CO molecules that reach the cluster via reverse spillover. Our calculations also show that repeated CO oxidation with the system $O_M Pd_9$/MgO(FC), M=3,2,1, is possible via reaction routes similar to the ones described above for M=4, suggesting a possible complete re-metallisation of the cluster. However, Pd_9/MgO(FC) strongly adsorbs CO molecules on top of Pd-Pd bridges with adsorption energies of 1.7 eV. This inhibits subsequent O_2 dissociation and poisons the catalyst. Hence, our results indicate that MgO-supported Pd clusters would catalyse CO oxidation best under O_2-rich conditions.

Our barriers for CO_2 formation (0.6-1.0 eV), are reasonable when compared to corresponding model catalyst experiments[3] with size-selected Pd_N/MgO(FC), where CO_2 was formed between 200-600 K.[34] Furthermore, the poisoning effect[34] observed for CO deposited prior to O_2 agrees well with our finding that pre-adsorption of CO inhibits subsequent O_2 dissociation.

Acknowledgments

We thank U. Heiz for the communication of unpublished results and fruitful discussions. This work was supported by the Deutsche Forschungsgemeinschaft, the Fraunhofer MAVO for Multiscale Materials Modelling (MMM) and the Academy of Finland (AF).

References

1. Chemisorption and Reactivity on Supported Clusters and Thin Films; Lambert, R. M., Pacchioni, G., Eds.; Kluwer: Dordrecht, 1997; p 395-424.

2. H. Grönbeck, Topics in Catalysis **28**, 59 (2004).
3. U. Heiz and W-D. Schneider, in *Metal Clusters at Surfaces*, edited by K.H. Meiwes-Broer (Springer, Berlin, 2000) p. 237-273 and references therein.
4. C. Becker and C.R. Henry, Surf. Sci. **352**, 457 (1996).
5. H.J. Freund, Adv. in Catal. **45**, 333 (2000).
6. S.H. Shaikhutdinov et al. *et al.*, Surf. Sci. **501**, 270 (2002).
7. J. Mizei et al., Thin Solid Films **391**, 209 (2001).
8. M. Westphalen et al., Sol. Ener. Mater. & Sol. Cells. **61**, 97 (2000).
9. S. Fukami, A. Ohno and A. Tanaka, Mater. Trans. **45**, 2012 (2004).
10. P. Mars and D.W. van Krevelen, Chem. Eng. Sci. **3**, 41 (1954).
11. B.L.M. Hendriksen et al., Surf. Sci. 552, 229-242 (2004).
12. Cluster of atoms and molecules, edited by H. Haberland (Springer, Berlin, 1994).
13. B. Hammer and J.K. Norskov, in Chemisorption and Reactivity on Supported Clusters and Thin Films; Lambert, R. M., Pacchioni, G., Eds.; Kluwer: Dordrecht, 1997; p 331-351.
14. M. Moseler at al., Phys. Rev. Lett, **86**, 2545 (2001).
15. M. Moseler, H. Häkkinen and U. Landman, Phys. Rev. Lett. **89**, 176103 (2002).
16. S. Abbet et al., Phys. Rev. Lett. **86**, 5950 (2001).
17. B. Huber et al., Nature Materials **5**, 44 (2006).
18. B. Huber and M. Moseler, EPJD online (2007).
19. W. Kohn and L.J. Sham, Phys. Rev. A **140**, 1133, (1965).
20. R. Parr and W. Yang, *Density functional theory of atoms and molecules* (Oxford university press, 1989).
21. N. Troullier, and J.L. Martins, Phys. Rev. B **43**, 1993 (1991).
22. J.P. Perdew et al., Phys. Rev. Lett. 77, 3865 (1996).
23. R. Barnett and U. Landman, Phys. Rev. B **48**, 2081 (1993).
24. G. Henkelman and H. Jonsson, J. Chem. Phys. **113**, 9978 (2000).
25. F.M. Richards, J. Mol. Biol. **82**, 1 (1974).
26. J. Tersoff and R. Hamann, Phys. Rev. Lett. **50**, 1998 (1983).
27. S. Abbet et al., J. Am. Chem. Soc. **123**, 6172 (2000).
28. A. Eichler, F. Mittendorfer and J. Hafner, Phys. Rev. B **62**, 4744 (2000).
29. B. Huber et al., Comp. Mat. Sci. **35**, 371 (2006).
30. E. Lundgren et al., Phys. Rev. Lett. **88**, 246103 (2002).
31. M. Todorova et al., Surface Science **541**, 101 (2003).
32. U. Heiz and E.L. Bullock, Mater. Chem. **14**, 564 (2004).
33. C.J. Zang and P. Hu, J. Am. Chem. Soc. **123**, 1166 (2001).
34. U. Heiz, A. Sanchez and S. Abbet, Chem. Phys. **262**, 189 (2000).

Structural, Electronic, and Optical Properties of Semiconductor Nanocrystallites

Luis E. Ramos and Friedhelm Bechstedt

Friedrich-Schiller-Universität Jena, Institut für Festkörpertheorie und -optik
Max-Wien-Platz 1, D-07743 Jena, Germany
E-mail: lramosde@yahoo.de,bechstedt@ifto.physik.uni-jena.de

We investigate the electronic and optical properties of large Si and Ge nanocrystallites by means of an *ab initio* pseudopotential method based on density-functional theory and on the GW approximation (Green function and screened Coulomb potential). The nanocrystallites studied may contain more than four hundred atoms and more than one thousand electrons. We find important effects of the oxidation, oxide capping, germanium (silicon) capping, and different surface terminations, and doping (group III, group-IV, and group-V impurities) on the electronic structure and optical absorption spectra are studied semiconductor nanocrystallites.

1 Introduction and Motivation

Silicon nanocrystallites (NCs) have been intensively investigated in the last years due to their interesting quantum confinement properties.[1] The strong spatial localization of electrons and holes in Si NCs can enhance radiative recombination rates and give rise to luminescence. Among other known applications, research on Si NCs could lead to optoelectronic devices compatible with the consolidated Si technology. Optical gain in Si NCs has been reported,[3] and new devices have recently been suggested.[4] The most accepted models to explain the PL in nanosized materials are related to quantum confinement effects.

The role of oxygen in the peak position and efficiency of the PL has been demonstrated for oxidized Si NCs.[5] While the PL blue shift from Si NCs embedded in SiO_2 matrices and from the oxidation of porous Si seems to support the quantum confinement model, there is evidence that a PL red shift is caused by defect states.[5] Our calculations consider some Si NCs whose surface is completely oxidized,[6,7] which can explain better the variation of properties of nanometer-sized Si dots with backbond oxidation in comparison to those for H and OH passivations.[8]

Electrical and mechanical properties resulting from the combination of group-IV compounds such as Si and Ge indicate that their heterostructures and nanostructures are very promising materials in device technology.[9] Direct applications of the SiGe technology comprehend high-frequency transistors, infrared photodetectors and photodiodes, and solar cells. Besides new features in the electronic properties, SiGe-based NCs can provide different quantum confinement for carriers and charge-storage mechanisms, the latter being useful for low-power memory devices.

The influence of doping on the electronic properties observed in porous Si (pSi) composed by Si NCs of different sizes is still being investigated.[10] Doping of Si NCs is expected to introduce additional levels close to the HOMO (the highest occupied molecular orbital) or LUMO (lowest unoccupied molecular orbital) in the same fashion as it does for Si bulk. This fact could enhance the intensity of absorption and emission spectra and be useful in some applications.

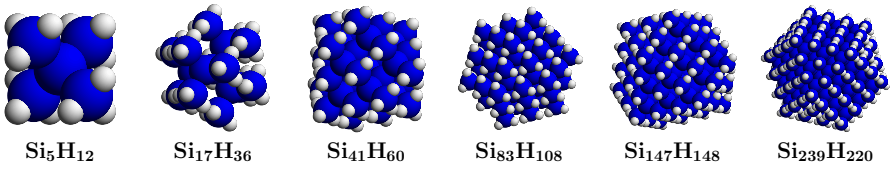


Figure 1. Faceted semiconductor nanocrystallites studied. The dark spheres correspond to Si or Ge atoms and the light ones to H atoms.

In order to model Si (Ge) NCs with a reasonable size and oxidation of Si NCs, a considerable computational power is needed, since a large number of electronic states is involved. In the following we describe the main results obtained for oxidation, capping, and doping of Si NCs. In Sec. 2 we describe the methods and approaches used in the calculations. The results are presented in Sec. 3 and in Sec. 4 we summarize our main results and conclusions.

2 Methods

We employ the density-functional theory (DFT), local-density approximation (LDA) or generalized-gradient approximation (GGA), as implemented in the Vienna *Ab initio* Simulation Package (VASP).[11,12] Projector-augmented-wave (PAW) data sets are applied to treat the interaction of electrons with the nuclei analogously to the pseudopotential method. The eigenfunctions and eigenvalues of the many-electron system are determined by solving the Kohn-Sham (KS) equations. The PAW data sets are suitable to calculate optical matrix elements with good accuracy.[13] VASP uses a plane-wave basis, the supercell approximation in the reciprocal space, and iterative matrix-diagonalization schemes, which are based on optimized minimization algorithms such as conjugate-gradient method, residual band-by-band minimization, and block Davidson.[14] The energy cutoffs required for the planewave basis are relatively low, which allows us to treat efficiently systems with hundreds of atoms. VASP requires standard linear-algebra libraries such as LAPACK (Linear Algebra PACKage) and BLAS (Basic Linear Algebra Subprograms),[15] but others such as ESSL (Engineering Scientific Subroutine Library), ScaLAPACK (Scalable LAPACK), ATLAS (Automatically Tuned Linear Algebra Software)[16] can also be used. The algorithms implemented in VASP allow several parts of the code to be parallelized using MPI (Message Passing Interface), e.g. parallelization over electronic bands and planewave coefficients.

The NCs have a shell-like structure and lie in large simple-cubic supercells. Starting from a central Si atom and assuming tetrahedral coordination, the NCs are constructed by adding the neighbouring atoms shell by shell. The dangling bonds at the NC surface are passivated with H atoms. In the case of doping, we consider additional Si NCs with a spherical-like shape. The faceted NCs investigated are shown in Fig. 1. To simplify the description of the NCs, in the following we denote them according to the number of atom shells in their core and in their capping shells: Si_5H_{12}, $Si_{17}H_{36}$, $Si_{41}H_{60}$, $Si_{83}H_{108}$, $Si_{147}H_{148}$, and $Si_{239}H_{220}$ are respectively Si2, Si3, Si4, Si5, Si6, and Si7. Oxidation in Si NCs is considered via the passivation of the dangling bonds with hydroxyls (OH) or as a

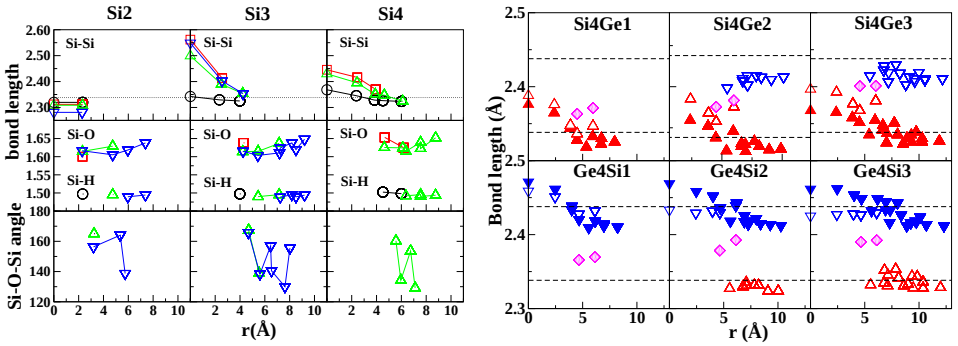


Figure 2. Left: Average bond lengths (in Å) and Si-O-Si angles (in degrees) versus distance from centre of the NCs for Si cores with varying size. The different oxidation stages are indicated by circles (H passivation), squares (OH passivation), up triangles (one oxide shell), and down triangles (two oxide shells). The dotted lines indicate the bulk Si bond length. Right: Si-Si bond lengths (up triangles), Ge-Ge bond lengths (down triangles), and Si-Ge bond lengths (diamonds) versus distance of the atom from the centre of the NC for the Ge-capped Si NCs (upper panels) and for the Si-capped Ge NCs (lower panels). The Si-Si bulk bond length (2.34 Å) and the Ge-Ge bulk bond length (2.44 Å) are indicated by the dashed lines. The filled triangles correspond to the Si and Ge NCs passivated with H and with the same number of shells as the capped Si (Ge) NCs in each frame.

backbond oxidation of one and two shells immediately below the surface.[7] To achieve a total energy minimum and negligible interatomic forces, we perform the ionic relaxation for all NCs without symmetry constraint. Spin polarization is taken into account in the case of doped Si NCs.

The optical matrix elements are calculated within the framework of the independent-particle approximation (DFT/KS) or independent-quasiparticle approximation (GW). The optical absorption spectra are identified as the imaginary part of the corresponding dielectric function and a Lorentzian broadening of 0.1 eV is applied in all spectra.

3 Results

3.1 Structural Properties

In Fig. 2 we summarize the structural properties of oxidized Si NCs and Si (Ge) capped NCs after ionic relaxation. As a result of ionic relaxation, the Si-Si bond lengths in the oxidized Si NCs deviate from the bulk value differently in the core and close to the surface of the NCs. The Si-Si bond lengths of very small Si NCs are slightly shortened with respect to the bulk value, whereas the largest cores Si3 and Si4 exhibit an expansion of the Si-Si bonds at the core. Oxidation of the Si NCs increases the expansion of the Si-Si bonds. The Si-O bonds lengths increase slightly close to the outermost shells of the NC, while the Si-H bond lengths practically do not vary versus the NC radii. As a result of the relaxation, the Si-O-Si angles tend to decrease versus the distance from the centre of the Si NC.[7]

Si and Ge NCs passivated with H also exhibit an expansion of the Si-Si or Ge-Ge bond lengths with respect to their bulk value at the centre of the NC and a contraction of the same bonds near the surface.[7] For the Ge-capped Si NCs, the Si-Si bonds lengths follow

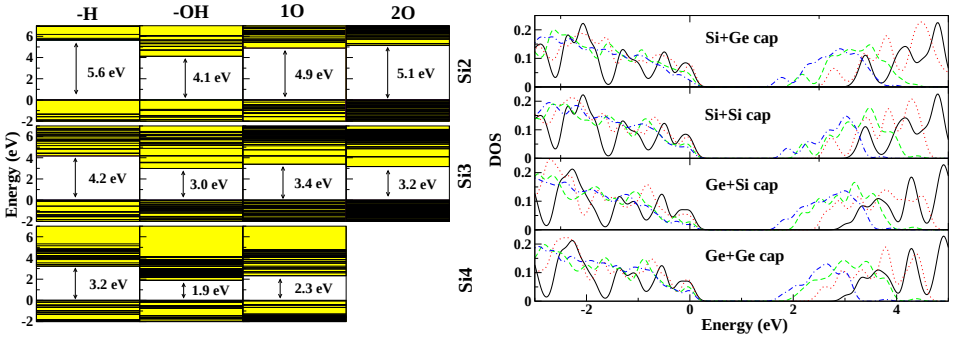


Figure 3. Left: Electronic levels of the Si NCs with different passivation and oxide coverage as calculated by DFT-LDA. In the columns, from the left to the right, the electronic levels for NCs passivated with H, passivated with OH, with one covering oxide shell, and with two covering oxide shells. The number of Si atoms in the core of the NCs is indicated at the right side. The values of the HOMO-LUMO gaps are indicated by the arrows. Right: Density of states for capped and uncapped NCs with 4 shells (solid lines, uncapped NC), 5 shells (dotted lines), 6 shells (dashed lines), and 7 shells (dot-dashed lines). The HOMO of all the structures is aligned to the energy zero.

the same trend as for the Si NCs. The Ge-Ge bond lengths in the capping shells are smaller than the one for Ge bulk and tend to have similar values with increasing number of capping shells. However, for the Si-capped Ge NCs there is a tendency for both Ge-Ge and Si-Si bond lengths to be slightly smaller than their corresponding bulk values and to vary much less than for the Ge NCs. The Si-Ge bonds at the interface between the capping shell and the NC core increase with an increasing size of the NC in both Si-capped and Ge-capped structures. The Si-Si and Ge-Ge bond lengths in the Si-capped Ge NCs indicate a small strain of the bonds which is even lower than the one for Ge NCs.

3.2 Electronic Properties

The effect of oxide coverage on the DFT-LDA KS eigenvalues of Si NCs is shown in Fig. 3. Both the passivation of the Si NC with OH and the oxide coverage give rise to a narrowing of the HOMO-LUMO energy gap. The narrowing of the gap is more pronounced in the case of passivation with OH than in the case of an oxide coverage of the Si NCs. As the size of the Si NC core increases, the HOMO-LUMO gap narrows as a consequence of the reduced spatial confinement of the corresponding states. In Fig. 3, we show the effect of the capping shells in Ge and Si NCs on the electronic density of states (DOS). As a rule, the increasing size of the NCs leads to a reduction of the quantum confinement and the HOMO-LUMO energy gap narrows. As shown in Fig. 3, the structures that contain more Ge atoms tend to have a narrower gap than the structures which contain more Si atoms. In this sense, our well-ordered capped NCs can be considered as one of the possible configurations of a SiGe random-alloy nanocrystallite. One can therefore identify two competing effects in determining the HOMO-LUMO gap of the capped Si and Ge NCs. One is due to quantum confinement and the other one is dictated by the overall composition of the NC. The reduction of quantum confinement combined with the increase of Ge contents is very effective in shortening the HOMO-LUMO gap of the Ge-capped Si NCs.

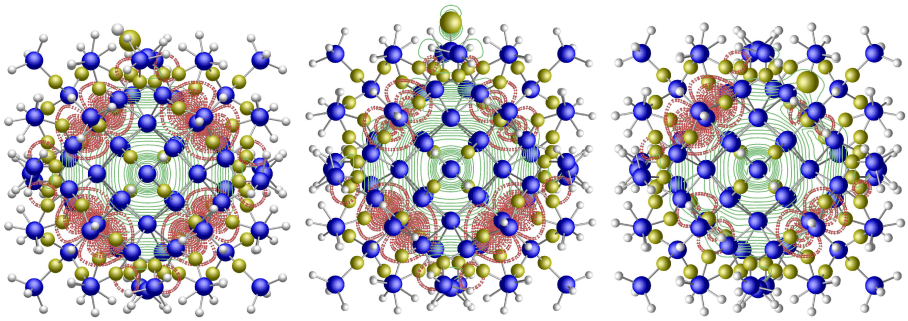


Figure 4. From the left to the right we show optimized structures for the single terminations -OH and =O, and the oxygen vacancy in the oxidized NC Si4-1O. The dashed red lines indicate the charge distribution of HOMO and the solid green lines indicate the charge distribution of the LUMO. The Si (blue spheres), O (yellow spheres), and H (white spheres) atoms in the Si NC are represented. The largest spheres in each picture correspond to the termination or site where the defect is localized.

3.3 Defects and Doping of Si Nanocrystallites

In Fig. 4 we show the effect of a modified surface termination and the presence of defects in oxidized Si NC. The model system is a Si NC with one oxide shell and passivated with H (Si4-1O). In order to study the effect of oxygen-related defects on Si NCs, we consider an oxygen vacancy V_O in a Si-O-Si bond, a OH replacing a H atom, and a double-bonded oxygen atom replacing two H atoms (Si=O) at the surface of an oxidized NC.[6] The V_O is introduced in a Si-O-Si bond, whereas the -OH and =O defects are located at the central atom of a facet. Practically no influence on the electronic structure (not shown) is observed when a H atom is replaced by a OH group. Slight changes occur for the probability distributions of HOMO and LUMO according to Fig. 4. While the HOMO is mostly localized at the interface between the core and the oxide shell of the NCs, the LUMO is delocalized and spreads over the whole NC extent.[6] Although the influence in the case of the =O and V_O defects is more pronounced than that of the -OH defect, the regions where the HOMO and LUMO are localized remain the same as for the case of the defect-free oxidized Si NC.

The formation energies of group-III (B and Al), group-IV (C and Ge), and group-V (N and P) impurities versus the reciprocal radii of the Si NCs are shown in Fig. 5 and calculated by $\Omega_{\rm f}[X] = E_{\rm tot}[X] - E_{\rm tot} + \mu_{\rm Si} - \mu_X$, where $E_{\rm tot}[X]$ (X=B, Al, C, Ge, N, and P) is the total energy of the doped Si NC, $E_{\rm tot}$ is the total energy of the undoped Si NC, and $\mu_{\rm Si}$ and μ_X are estimates for the chemical potentials of the reservoirs for the Si and for the X species, respectively.[10] The influence of shape and size of the NCs on the impurity formation energies is shown in Fig. 5 for doped Si NCs versus their reciprocal radii. The formation energies of Ge and P impurities depend mainly on the size of the particle and have an almost linear dependence on the reciprocal radius of the Si NC. While the formation energy of the first-row elements C and N may increase for large Si NCs, the formation energy of B, Al and P decreases. Since B leads to symmetry lowering as Al does and large deformation as P does, the impurity formation energies of B are high in small Si NCs. Only the formation energy of the Ge impurity is nearly independent of the size of the nanocrystallite. This fact is related to the similar chemical properties of Si and

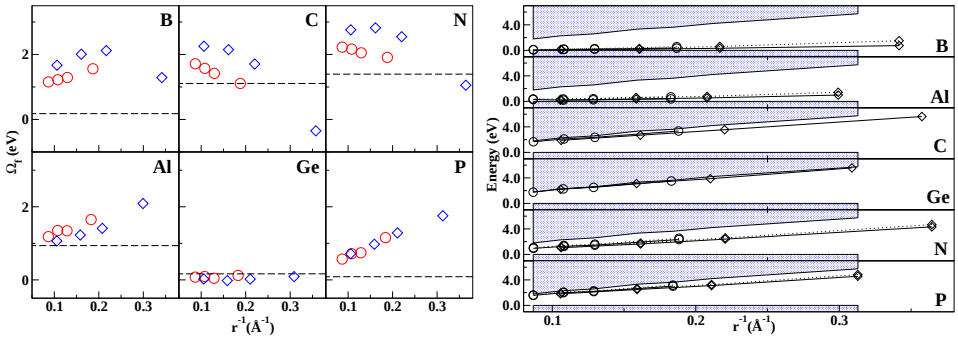


Figure 5. Left: Impurity formation energies of B, Al, C, Ge, N, and P versus the reciprocal radii of the doped Si NCs. The energy values are given for faceted (diamonds, see Fig. 1) and spherical-like (circles) doped Si NCs. The dashed lines indicate the impurity formation energies in bulk Si. Right: Energetic position of the KS impurity levels versus the reciprocal radii of the spherical-like (circles) and faceted (diamonds) Si NCs. The solid and dotted lines correspond to the spin-up and spin-down impurity levels. For the C- and Ge-doped Si NCs, the position of the LUMO is shown instead of the impurity level.

Ge. A comparison between faceted and spherical-like NCs shows that the incorporation of the impurity in Si NCs is more favourable, if the NC has facets in the case of Al, nearly shape-independent for Ge and P, and costs more energy for the other impurities considered. Except for Ge and the C and N doping of the smallest Si NCs, the impurity formation energies of the Si NCs are higher than in Si bulk. This can be explained in terms of longer Si-X bond lengths and the symmetry lowering of the doped Si NCs, which require additional relaxation of the atomic positions in comparison with bulk Si.

In Fig. 5 the position of the KS impurity levels in the energy-gap region as well as the HOMO-LUMO gap of the undoped Si NCs shows nearly a linear dependence versus the reciprocal radii of the NCs. An increase of quantum confinement leads to deep impurity levels in the doped Si NCs. In Si NCs the B and Al impurities induce shallow acceptor levels, whereas the N and P impurities induce deep and shallow donor levels, respectively. Only small variations of the HOMO-LUMO gaps due to the shape of the Si NCs can be perceived in figure 5 and do not vary significantly with the shape of the Si NCs.

3.4 Optical Absorption Spectra

In Fig. 6, the optical absorption spectra within the independent-quasiparticle approximation (GW) and HOMO-LUMO gaps are presented for different Si NC core sizes, passivations, and capping oxide shells. Si NCs passivated with hydroxyls have a more pronounced narrowing of the gap than the ones with an oxide coverage and the ones passivated with H. The most intense optical transitions occur at higher energies than the HOMO-LUMO gap and tend to increase with increasing oxygen contents and increasing size of the Si NCs. The indirect-gap character of bulk Si is transferred to all defect-free Si NCs independently of how oxygen is incorporated, via passivation with OH or via backbond oxidation. Our results indicate that the main influence of oxygen on the optical properties of defect-free Si NCs are the shifts of onset of the absorption in the case of the backbond oxidation and

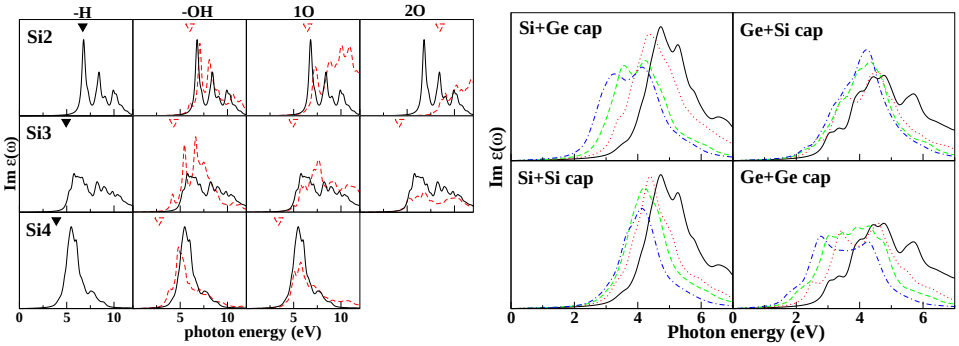


Figure 6. Left: Optical absorption spectra (arbitrary units) within the independent-quasiparticle approximation for Si NCs passivated with H (solid lines), passivated with OH, with one oxide shell 1O, and with two oxide shells 2O (columns), (dashed lines) and for Si2, Si3, and Si4 cores (rows). The HOMO-LUMO gaps are indicated by triangles (solid/dashed). Right: Optical absorption spectra of NCs with Si core (left panels) and NCs with Ge core (right panels). The upper panels correspond to the capped structures, whereas the lower panels show the spectra of Si and Ge NCs. For Si4(Ge4) NC core the spectra are shown for the case without capping shells (solid lines), one (dotted lines), two (dashed lines), and three (dot-dashed lines) capping shells.

passivation with OH. Backbond oxidation blue shifts the onsets of absorption of Si NCs with oxygen contents of bulk Si. The larger the Si NCs with backbond oxidation are, the smaller are the relative blue shifts of the onsets of absorption.

In Fig. 6, we present the optical absorption spectra of the capped and uncapped Si(Ge) NCs. As for the DOS in Fig. 3, the absorption spectra in Fig. 6 show clearly the importance of quantum confinement in determining the HOMO-LUMO gap. The addition of Ge capping shells to both Si and Ge NCs shifts the peaks of the absorption to red. Conversely to what is depicted from Fig. 3, the absorption spectra of the Si-capped Ge NCs shows a compensation effect between the composition of the NCs and their quantum confinement properties given by their size. In the Si-capped Ge NCs the addition of Si atoms tends to increase the value of the HOMO-LUMO gap due to the composition, whereas the increasing number of shells and consequently the size of the NC tends to reduce its value. As a result, the absorption peaks are not so shifted to red as in the case of uncapped Ge NCs with the increasing of Si capping shells.

4 Summary and Conclusions

By means of an *ab initio* pseudopotential method based on density-functional theory we investigated large semiconductor nanocrystallites concerning their structural, electronic and optical properties, when they are oxidized, capped, and doped with impurities of the groups III, IV, and V. The overall effect of capping with oxide shells in Si NCs is to narrow the HOMO-LUMO gap and expand the Si-Si bonds at the core of the NC, in particular when the Si NCs are oxidized via passivation with OH. Oxygen-related defects in oxidized Si NCs may have influence on the spatial distribution of the HOMO and LUMO, though their influence on the electronic properties is small. Capping of Si NCs with Ge shells apparently leads to more strain than the capping of Ge NCs with Si shells. The onsets

of absorption of Si (Ge) NCs capped with Ge (Si) shells shifts due to the contribution of quantum confinement effects and overall composition of the capped NC. Except for Ge, the formation energy of an impurity in doped Si nanocrystallites tends to be higher than in bulk Si. The impurity formation energies depend significantly on the shape of the Si NCs. As for bulk Si, B and Al introduce shallow acceptor levels and N and P introduce donor levels in Si NCs. There is less dependence of the impurity energy levels with respect to the shape of the Si NCs. Doping with C may narrow the HOMO-LUMO gap, whereas doping with Ge practically does not change the gap. Our results help to clarify some aspects of electronic structure and optical spectra of semiconductor nanocrystallites.

References

1. A.G. Cullis, L.T. Canham, and P.D.J. Calcott, *The structural and luminescence properties of porous silicon*, J. Appl. Phys. **82**, 909–965, 1997.
2. O. Bisi, S. Ossicini, and L. Pavesi, *Porous silicon: a quantum sponge structure for silicon based optoelectronics*, Surf. Sci. Rep. **38**, 1–126, 2000.
3. L. Pavesi, L. Dal Negro, C. Mazzoleni, G. Franzò, and F. Priolo, *Optical gain in silicon nanocrystals*, Nature (London) **408**, 440–444, 2000.
4. J. Linnros, *Nanocrystals brighten transistors*, Nature Materials (London) **4**, 117–119, 2005.
5. M.V. Wolkin, J. Jorne, P.M. Fauchet, G. Allan, and C. Delerue, *Electronic States and Luminescence in Porous Silicon Quantum Dots: The Role of Oxygen*, Phys. Rev. Lett. **82**, 197–200, 1999.
6. L.E. Ramos, J. Furthmüller, and F. Bechstedt, *Reduced influence of defects on oxidized Si nanocrystallites*, Phys. Rev. B **71**, 035328-1–035328-7, 2005.
7. L.E. Ramos, J. Furthmüller, and F. Bechstedt, *Effect of backbond oxidation on silicon nanocrystallites*, Phys. Rev. B **70**, 033311-1–033311-4, 2004.
8. L.E. Ramos, J. Furthmüller, and F. Bechstedt, *Influence of Oxygen on Optical Properties of Si Nanocrystallites*, Appl. Phys. Lett. **87**, 143113-1–143113-3, 2005.
9. L.E. Ramos, J. Furthmüller, and F. Bechstedt, *Quantum confinement in Si- and Ge-capped nanocrystallites*, Phys. Rev. B **72**, 045351-1–045351-8, 2005.
10. L.E. Ramos, E. Degoli, G. Cantele, S. Ossicini, D. Ninno, J. Furthmüller, and F. Bechstedt, *Structural features and electronic properties of group-III, group-IV, and group-V doped Si nanocrystallites*, J. Phys. C: Condens. Matter. **19**, 466211-1–466211-12, 2007.
11. G. Kresse and J. Furthmüller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*, Comput. Mater. Sci. **6**, 15–50, 1996.
12. G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculation using a plane-wave basis set*, Phys. Rev. B **54**, 11169–11186, 1996 .
13. G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*, Phys. Rev. B **59**, 1758–1775, 1999 .
14. `http://cms.mpi.univie.ac.at/vasp/`.
15. `http://www.netlib.org`.
16. `http://math-atlas.sourceforge.net`.

Condensed Matter: From the Schrödinger Equation of Many Electrons and Nuclei to Fluid Flow

Kurt Binder[1] **and Alejandro Muramatsu**[2]

[1] Institut für Physik, Johannes-Gutenberg-Universität
55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

[2] Institut für Theoretische Physik III, Universität Stuttgart
70550 Stuttgart, Germany
E-mail: mu@theo3.physik.uni-stuttgart.de

In principle, any phenomena in condensed matter could be theoretically understood if one were able to solve the many-body Schrödinger equation of electrons and nuclei (interacting with just the appropriate Coulomb potentials) and use this solution as input to the formalism of (quantum) statistical mechanics. Unlike elementary particle physics, where the nature of some of the basic interactions are still debated, the starting point of the theory of condensed matter is perfectly well defined. However, the problem posed is far beyond reach even of the most powerful computers, also in the foreseeable future, and hence a large variety of approximations is needed, where (i) the Hamiltonian is reduced to some model Hamiltonian (often involving some unknown phenomenological parameters), and the type of model depends very much on the systems one wants to describe, and (ii) implicit approximations are necessary in the simulation methods used to study these models. In fact, depending on the problems that are studied a wide variety of models and simulation methods are encountered, and this fact will become apparent from the articles that will follow in this section. Of course, these articles constitute only a small selection of problems in the science of condensed matter that are studied using the resources of NIC; in addition, we note that many articles that appear in the sections on chemistry, materials science, polymers and soft matter also deal with condensed matter, of course.

Now progress in the simulation studies of condensed matter systems is gained by improvements of algorithms on the one hand, and by better computer hardware (faster processors, larger memory, etc.) on the other hand. In favourable cases, these two aspects combine and allow to deal with problems that were intractable so far.

An example for this statement is the article by Dolfen and Koch on the attempt of a realistic description of the quasi-one-dimensional organic conductor TTF-TCNQ. This system can be modelled as a strongly correlated organic metal, and hence a description in terms of a (slightly extended) Hubbard model results. The appropriate parameters of this model are estimated by all-electron density functional (DFT) methods. The hamiltonian then is diagonalized applying the Lanczos algorithm, to obtain the ground state energy and excitations. However, only finite clusters can be dealt with exactly, to deal with the infinite chain hopping between clusters is treated by a “cluster perturbation theory”, which allows to study fascinating physics like spin-charge separation in the excitation spectrum. It should be emphasized that due to excessive memory requirements this approach was

only feasible using a massively parallel implementation at the Blue Gene/L system JUBL.

Dealing with a strongly correlated electron systems is also the focus of the article by Martin et al. on the numerical simulation of heavy fermion systems. A generic model for such systems is the Kondo lattice model, which is tackled here by a dynamical cluster approximation (a finite sized cluster is embedded in a bath of the remaining electrons), using a quantum Monte Carlo algorithm. Again use of a highly parallelized implementation on high-performance computers is crucial for a success of this study, which allowed first steps towards a study of the magnetic order-disorder transition of this model.

An interesting step on the (long) way towards the solution of the many-body Schrödinger equation is also the *Ab Initio* Molecular Dynamics (AIMD) method, applied in the chapter by Groß to study adsorption of hydrogen molecules on palladium surfaces. Using density functional theory within the generalized gradient approximation, the substrate surface at $T = 0K$ was prepared, and then at least 200 microcanonical Molecular Dynamics (MD) trajectories of the adsorption and desorption of the hydrogen molecules were generated, for each choice of the initial kinetic energy of the molecule. The simulations again need an efficiently parallelized code and unravel the crucial role of the substrate degrees of freedom (clean and hydrogen-precovered surfaces are studied) on the dissociative adsorption process.

When one studies large systems at high temperatures, however, the above types of techniques are still unfeasible, and one must rely on phenomenological potentials that are used in classical MD methods. However, if these potentials are carefully chosen, a reasonably good description of real systems is possible. This fact is demonstrated in the article by Kerrache et al. on the simulation of multicomponent liquids. Two systems are studied: molten silica and AlNi alloys, both in the crystalline state and in the melt. For silica, a new and more accurate potential is defined, from a new fitting scheme to Car Parrinello Molecular Dynamics (CPMD). Like AIMD, the starting point is also DFT, but the dynamics is stochastic rather than microcanonical. For the metallic alloy, a standard embedded-atom potential from the literature is used, but the new methodic aspect in this part of the study is the very accurate estimation of the melting temperature from recording the temperature dependence of the crystal-melt interface velocity. Due to the necessary sample averaging the use of a high performance computer such as JUMP also was indispensable for the success of this project.

While such a chemically realistic description of the dynamics of fluids on the atomistic scale requires MD methods, coarse-grained models of fluids can also be studied by Brownian Dynamics (BD) or Lattice Boltzmann (LB) methods. The latter approach is implemented in the article by Harting and Kunert, where flow in thin channels with rough surfaces is studied, to find out how the slip length depends on the wall roughness. The BD method is applied by Bürzle et al. for the study of transport phenomena of colloids in microchannels: the colloidal particles move as a regular array of several rows, but also transitions from n to $n - 1$ in the number of "lanes" for this motion are observed. Other coarse-grained models concern the phase behaviour of colloidal particles in the periodic potential by laser fields (which is studied by Monte Carlo (MC)) and the structure of lipid bilayers with incorporated proteins (a model for a biological membrane) in solution. Due to the large number of particles ("only" 4800 lipid molecules but 90 000 solvent particles) the need for (geometric) parallelization on a high-performance computer such as JUMP again is clearly apparent.

Advanced MC methods are also the technique of choice in the article by Bittner et al. on free-energy barriers of spin glasses (Ising spins interact with random couplings, assuming either infinite range (SK model) or nearest neighbour forces (EA model)). Due to the extremely slow relaxation and the fact that the free energy barriers are not self-averaging, this problem is by no means simple. Progress is obtained by combining the "multi-overlap algorithm" with "parallel tempering"; thus again the need for a parallel implementation on a high-performance computer such as JUMP immediately becomes apparent. In this way, information can be extracted on how the barrier heights scale with the number of spins in the system.

Numerical Simulations of Heavy Fermion Systems

Lee C. Martin, Fakher F. Assaad, and Werner Hanke

Institut für Theoretische Physik und Astrophysik, Universität Würzburg
Am Hubland, 97074, Würzburg, Germany
E-mail: {Lee.Martin, Fakher.Assaad, Werner.Hanke}@physik.uni-wuerzburg.de

The Kondo lattice model describes a lattice of magnetic impurities embedded in a metallic host and is an appropriate starting point for the understanding of so-called heavy fermion systems. Those systems show a plethora of competing phases: a metallic Fermi-liquid state with effective mass exceeding by order of magnitudes the bare electron mass, magnetically ordered states as well as the coexistence of superconductivity and magnetism. Quantum phase transitions between the above mentioned phases as well as the mechanism triggering superconductivity remain unresolved issues of great interest in solid state physics. With the use of large scale numerical simulations in the framework of the Dynamical Cluster Approximation, we have investigated the evolution of the Fermi surface across the magnetic order-disorder transition. This is a central issue which lies at the heart of a theoretical understanding of this transition. Here we summarize our results which show a change in the topology of the Fermi surface across the transition thus supporting recent Hall effect experiments.

1 Introduction

In the 1960's the study of localized magnetic moments was placed firmly in the emerging field of strongly correlated electron systems when Anderson first identified interactions between localized electrons as the driving force for local moment formation[1]. Early experimentalists realized that local moment formation on magnetic iron ions dissolved in non-magnetic metals is dependent on the host[2,3]. The magnetic susceptibility of a system of iron dissolved in a niobium-molybdenum alloy was seen to follow a Curie-Weiss law for compositions close to molybdenum, indicating the presence of local moments. The low temperature limit of Curie-Weiss susceptibility is given by the Kondo temperature, the temperature at which the local moment is screened due to the formation of an entangled spin singlet state of the local moment and surrounding conduction electrons. Kondo physics lies in direct analogy to the phenomena of asymptotic freedom that governs quark physics. Like the quark, at high energies the local moment in the metallic host is free, but at energies below the Kondo temperature, it interacts so strongly with the surrounding electrons that it becomes screened or confined. The physics of local moment screening, manifests itself in a variety of properties of correlated electron systems including the Kondo resistivity minimum and also the formation of heavy fermion metals. With the latter, the presence of local moments greatly changes the metals properties with quasi-particles developing which may have an effective mass many times larger than the bare electron mass whilst still behaving as a Fermi liquid at low temperatures.

Such behaviour led Doniach to propose that the huge mass renormalization has its roots in a lattice version of the Kondo effect and that such heavy fermion systems should be modelled by the Kondo lattice model (KLM)[4]. The mass renormalization can be attributed to the coherent superposition of individual Kondo screening clouds and the resulting metallic state is characterized by a Fermi surface with Luttinger volume (Fermi surface volume)

containing both conduction and localized electrons. In its simplest form, the KLM describes a lattice of spin 1/2 magnetic moments coupled antiferromagnetically via an exchange coupling J to a single band of conduction electrons and is believed to capture the physics of heavy fermion materials such as $CeCu_6$.

Although the physics of the single impurity Kondo problem is well understood[5] the KLM still poses a problem half a century since its original conception. The difficulty with the lattice problem arises due to the presence of two competing energy scales. The first energy scale, given by the Kondo temperature, is associated with the screening of impurity spins via the Kondo effect. However, in the lattice problem, polarization of the conduction electron spins around a first magnetic impurity can couple to a second impurity leading to an effective interaction between impurity spins, the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction[6–8], and an associated second energy scale. This RKKY scale dominates at low values of the exchange coupling and is the driving force for the observed magnetic order-disorder quantum phase transitions in heavy fermion materials. The nature of this phase transition is of current interest following experimental results suggesting a sudden change in the Fermi surface topology at the quantum critical point (QCP) for the heavy fermion metal $YbRh_2Si_2$[9]. Tuning this system from the non-magnetic heavy fermion metallic phase to the antiferromagnetic metallic phase causes a rapid change in the low temperature Hall coefficient which is extrapolated to a sudden jump at $T = 0$. Since the low-temperature Hall coefficient is related to the Fermi surface volume the results are interpreted as showing a sudden reordering of the Fermi surface at the QCP from a *large* Fermi surface, where the local moment impurity spins are included in the Luttinger volume, to a *small* Fermi surface where the impurity spins drop out of the Fermi surface volume.

This issue forms one of the central issues of our work. In this article we briefly present this and other recent results of our numerical simulations of the KLM and attempt to highlight the essential need for high performance computing in this project.

2 Numerical Solution of the Kondo Lattice Model

We take the KLM in two dimensions as our model heavy fermion system with Hamiltonian given by

$$H = \sum_{\mathbf{k},\sigma} \epsilon(\mathbf{k}) c^\dagger_{\mathbf{k},\sigma} c_{\mathbf{k},\sigma} + J \sum_{\mathbf{i}} \mathbf{S}^c_{\mathbf{i}} \cdot \mathbf{S}^f_{\mathbf{i}} \tag{1}$$

with $c^\dagger_{\mathbf{k},\sigma}$ creating a conduction electron on an extended orbital with wave vector $\mathbf{k}$ and a z-component of spin $\sigma = \uparrow, \downarrow$. The spin $1/2$ degrees of freedom, coupled via J, are represented with the aid of the Pauli spin matrices $\boldsymbol{\sigma}$ by $\mathbf{S}^c_{\mathbf{i}} = \frac{1}{2} \sum_{s,s'} c^\dagger_{\mathbf{i},s} \boldsymbol{\sigma}_{s,s'} c_{\mathbf{i},s'}$ or the equivalent definition for $\mathbf{S}^f_{\mathbf{i}}$ using the localized orbital creation operators $f^\dagger_{\mathbf{i},\sigma}$. The KLM forbids charge fluctuations on the f-orbitals and as such the constraint of one electron per localized orbital must be included.

A numerical approach to solving the model is advantageous because such an approach will be non-biased. Due to the sheer size of the electron configuration (Hilbert) space of even a moderately small lattice of interacting electrons, exact solution methods, for example exact diagonalization, become impossible in practice. The quantum Monte Carlo

(QMC) approach stochastically samples the configuration space according to the statistical weight of a given configuration to arrive at an averaged solution with a statistical error dependant on the number of samples taken. However a common problem of QMC based simulations manifests itself as the minus sign problem in which the statistical weight of a sampled configuration may be negative. For a given algorithm the problem becomes severe if the average of this sign becomes small since then the number of stochastic samples required to give decent statistics increases exponentially. Previous QMC calculations of the KLM have been limited to finite sized lattices and in particular to particle-hole symmetry (half-filling) due to a severe minus sign problem.

In our project we examine the KLM in the hole-doped regime. We do this by employing the dynamical cluster approximation (DCA) in combination with a QMC algorithm for solving the model on the cluster.

3 The Dynamical Cluster Approximation with a Quantum Monte Carlo Cluster Solver

Within the DCA the original N-site lattice problem is approximated by a finite sized cluster of N_c sites embedded in a bath of the remaining electrons[10,11]. Interactions within the cluster are calculated exactly, whereas interactions between the cluster and the bath are accounted for at a mean field level in space coordinates whilst retaining the full dynamics in imaginary time. The approximation will be particularly valid for systems in which spatial fluctuations are short-ranged. This equates to assuming only a weak momentum dependency of correlation functions. In momentum space the Brillouin zone is divided into N_c patches and strict momentum conservation is relaxed by only requiring conservation for transfers taking the $\mathbf{k}$-vector out of the patch.

Our implementation of the DCA does not suppress the development of antiferromagnetic order as is the case with a standard DCA approach. Broken symmetry is taken into consideration by defining a unit cell of two lattice sites containing, in total, two conduction electrons and two localized f-electron orbitals. This is the smallest cluster, which we denote by $N_c^{AF} = 1$, with which we can capture antiferromagnetic ordering and is the smallest cluster we consider in this article. Fig. 1 defines the basis vectors used, the unit cell and site indices within a cell. Allowing for antiferromagnetic ordering, translational symmetry is now only assumed for the new basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ so that, in momentum space, the Brillouin zone is reduced to the magnetic Brillouin zone (MBZ). Patching of the MBZ is demonstrated in Fig. 1 for a cluster of size $N_c^{AF} = 4$. Importantly, the DCA self-consistent equation for the lattice Green function becomes a matrix equation since each site of the lattice defined by $\mathbf{a}_1$ and $\mathbf{a}_2$ is now a unit cell of four orbitals.

A basic requirement for using the DCA is the ability to effectively calculate the quantity $\Sigma^c[\bar{G}(\mathbf{K}, i\omega_n)]$, the self-energy on the cluster. In our work we achieve this via the QMC Hirsch-Fye impurity algorithm[12] proceeding as in Ref.[13].

From a technical point of view the bulk of the numerical effort lies within the calculation of the self-energy on the cluster via the QMC impurity algorithm. Here the computational time scales with $(\beta N_c^{AF})^3$ where β is the inverse temperature. Even for a *small* cluster containing 16 orbitals the computational effort required to obtain decent statistics is immense. High performance computing is at this stage essential and, since QMC methods are easily parallelized, highly applicable. Additionally, we note that during a typical

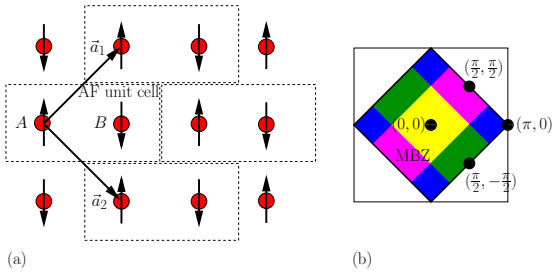


Figure 1. (a) Definition of the real space basis vectors, unit cell and orbital indices (b) The reduced Brillouin zone and k-space patching for a cluster with $N_c^{AF} = 4$

program run 90% of the time within the QMC algorithm is devoted to the calculation of outer products which may be achieved by using a BLAS library routine. The algorithm may therefore be highly optimized.

4 Results

As an initial test of our method we have carried out simulations at the particle-hole symmetric point ($t'/t = 0$ and $\langle n_c \rangle = 1$) and compared the results with previous lattice QMC simulations[13].

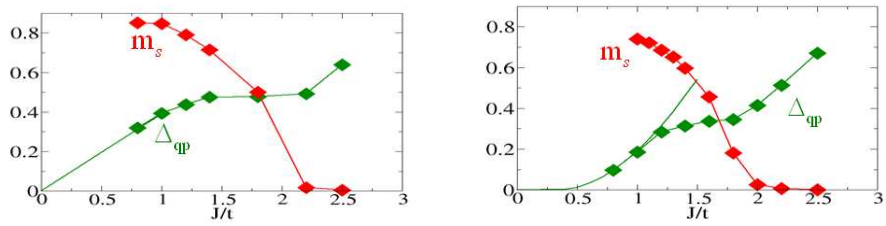


Figure 2. Left: DCA results for the staggered magnetization m_s and quasi-particle gap Δ_{qp} of the KLM at half-filling. Right: DCA results for the same quantities but now the model includes a next-nearest neighbour hopping term $t'/t = -0.3$

As shown in Fig. 2 (left side plot) the result for the staggered magnetization captures the phase transition from a paramagnetic to an antiferromagnetic phase with decreasing J/t. In addition, for small J/t the linear dependency of the quasi-particle gap, which appears in the results of Ref.[13], is well reproduced with the DCA results. At half-filling the Fermi-surface exhibits perfect nesting, so to examine the influence of this on the magnetization and quasi-particle gap we deformed the Fermi-surface by introducing a frustration term into the Hamiltonian in the form of a nearest-neighbour hopping $t'/t = -0.3$. The results (Fig. 2, right) still demonstrate a magnetically ordered phase, now at lower critical J/t, but the linear behaviour of the quasi-particle gap is lost for small J/t and the gap now appears

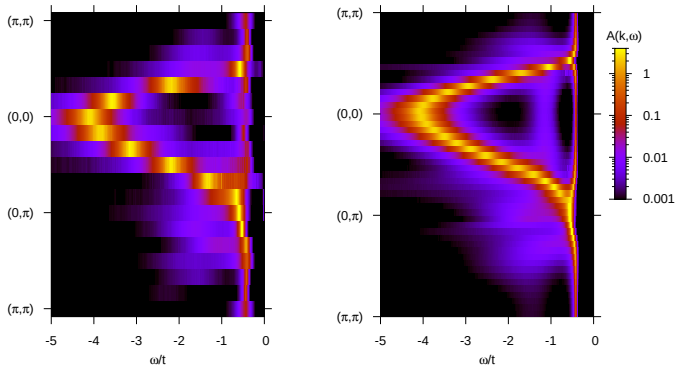


Figure 3. The single particle spectral function of the KLM at half-filling ($t'/t = 0$ and $\langle n_c \rangle = 1$) and with coupling $J/t = 1.2$ resulting from (left) a 12×12-lattice QMC simulation using the projective auxiliary field algorithm of Ref.[13] and (right) the DCA with cluster size $N_c^{AF} = 1$ and $\beta t = 40.0$

to follow a Kondo-scale, $\Delta_{qp}/t \propto e^{-W/J}$. This is an important result which concludes that a linear dependency of the quasi-particle gap is not generic to the system but rather is a direct consequence of particle-hole symmetry.

Initially remaining at half-filling we are able to show, by comparison of our band structure results with those of previous QMC calculations, that already a two-site cluster captures the key elements (Fig. 3). The parameters chosen for the plots are for an antiferromagnetic insulating point in the magnetic phase diagram. Although antiferromagnetic, the spectrum still retains a very flat heavy fermion band. This band, however, is then backfolded according to the symmetry effects introduced by the reduction of the Brillouin zone to a magnetic Brillouin zone as a direct result of antiferromagnetic ordering and breaking of the original translational symmetry. This behaviour is already well documented in the half-filled KLM, and the agreement of our results with those of previous studies confirms our belief that our DCA variant contains the necessary ingredients to investigate the delicate interplay between the RKKY interaction and Kondo physics.

In Fig. 4 we map out the ground state magnetic phase diagram of the KLM as a function of coupling J/t and conduction band hole-doping. Since we are interested in ground state properties the choice for the inverse temperature βt must be large enough to ensure we are below the smallest scale in the problem which will either be the coherence temperature or the RKKY scale. This requirement limits our simulations to the region $J/t \geq 0.8$ since the coherence scale decays exponentially with J/t and the computational time required by the QMC cluster solver scales as $(\beta N_c^{AF})^3$. Quite generally the onset of magnetism at small values of J/t is expected because the RKKY scale, set by $J^2\chi(\mathbf{q}, \omega = 0)$, then dominates over the Kondo scale given by $T_K \sim e^{-t/J}$. We note that the sign problem which restricts pure QMC simulations to the particle-hole symmetric case is not severe within the DCA framework.

Having confirmed the existence of an antiferromagnetically ordered metallic phase we now address the issue of the Fermi surface topology. In Fig. 5 we show results for the single particle energy excitation spectrum for parameter points with constant coupling J/t but on either side of the quantum phase transition. The energies ω/t are plotted relative to

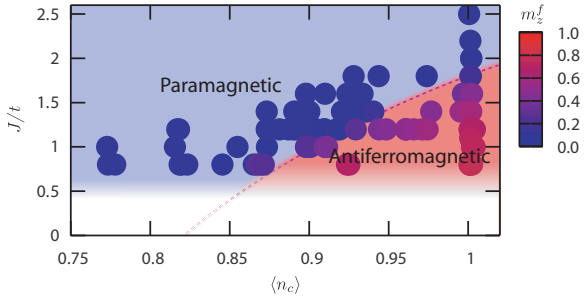


Figure 4. The magnetic phase diagram of the hole-doped KLM showing simulation results for the staggered magnetization m_z^f (colour-coded circles) as a function of coupling J/t and conduction electron occupancy $\langle n_c \rangle$. The shading of the antiferromagnetic and paramagnetic regions is intended only as a guide to the eye. We have included $t'/t = -0.3$ and the calculations are carried out with the $N_c^{AF} = 1$ cluster.

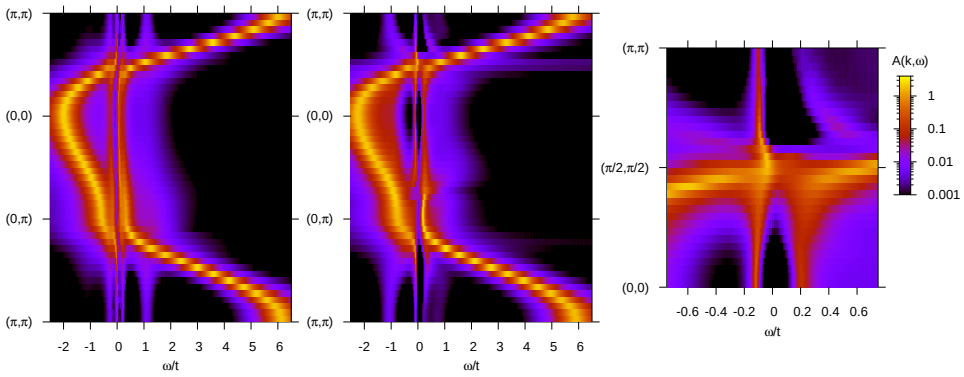


Figure 5. Left: DCA single particle spectrum for a simulation in the paramagnetic region ($J/t = 1.0$, $\langle n_c \rangle = 0.855$ and $\beta t = 40.0$). Centre: The spectrum of an antiferromagnetic point ($N_c^{AF} = 4$, $J/t = 1.0$, $\langle n_c \rangle = 0.977$ and $\beta t = 40.0$) across the phase transition but at the same coupling J/t. Right: Zoomed section of the central plot along the path $(0, 0)$ to (π, π) for energies around the Fermi energy

the Fermi energy and by observing where bands cross the Fermi energy we may deduce the topology of the Fermi surface. Beginning with the paramagnetic point the presence of very flat bands at the Fermi energy and around (π, π), and therefore a large effective electron mass, is no surprise since this region of parameter space can be well understood by considering a simple large-N mean field approach. The Fermi surface in this case is given by the left hand plot in Fig. 6 with unoccupied states around (π, π) and equivalent points. In the antiferromagnetic case we note the continued existence of flat heavy fermion bands but also that these bands drop below the Fermi energy around (π, π) and give way instead to pockets of unoccupied states around $(\pi/2, \pi/2)$. The resultant Fermi surface is represented in the right hand plot of Fig. 6. Evidently, the development of antiferromagnetic order is related to a complete reordering of the Fermi surface. For a more detailed analysis of such issues as whether the Luttinger volume counts the impurity orbitals in each case and other

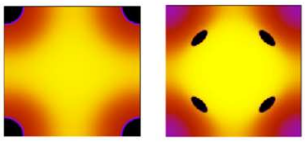


Figure 6. Depiction of the Fermi surface in the paramagnetic metallic state (left) and the antiferromagnetic metallic state (right)

related issues we refer the interested reader to our recent paper[14].

5 Conclusions

In this article we have reported on our current progress towards understanding the hole-doped, two dimensional KLM ground state. The DCA approach is able to capture the delicate interplay between Kondo screening and magnetic ordering, as demonstrated by comparison with earlier finite lattice QMC data at half-filling but the DCA has the clear advantage in the doped region since the sign problem is not severe. The central result presented here is the change of topology in the Fermi surface — large Fermi surface to hole pockets — when the system is tuned through the magnetic transition. This result has profound implications on the theoretical understanding of the phase transition. It invalidates the generic Hertz-Millis approach which relies of a Fermi surface topology which remains unchanged through the transition[15,16], and calls for alternative descriptions[17,18]. Clearly to pin down this important aspect, a more detailed investigation of the nature of the phase transition as well as the transition from one Fermi surface topology to the other in the region of the phase transition is required and ongoing. In parallel to this we are very interested in the effect of temperature on the single particle spectrum and work is also currently in progress to follow the evolution of the single particle excitation spectrum through different energy scales.

The statistically high quality data required to produce our current results and to achieve the above goals in the future, represents a significant investment in terms of computational cost. This ambitious project, which may be seen as a *grand challenge* in the field of correlated electron systems, would not be possible outside of the realms of high performance computing.

Acknowledgments

We would like to thank the Forschungszentrum Jülich for generous allocation of CPU time on the IBM Blue Gene/L and the DFG for financial support. We thank K. Beach, S. Capponi, S. Hochkeppel, T. C. Lang, T. Pruschke and M. Vojta for conversations.

References

1. P. W. Anderson, *Localized Magnetic States in Metals*, Phys. Rev., **124**, no. 1, 41–53, Oct 1961.

2. B. T. Matthias, M. Peter, H. J. Williams, A. M. Clogston, E. Corenzwit, and R. C. Sherwood, *Magnetic Moment of Transition Metal Atoms in Dilute Solution and Their Effect on Superconducting Transition Temperature*, Phys. Rev. Lett., **5**, no. 12, 542–544, Dec 1960.
3. A. M. Clogston, B. T. Matthias, M. Peter, H. J. Williams, E. Corenzwit, and R. C. Sherwood, *Local Magnetic Moment Associated with an Iron Atom Dissolved in Various Transition Metal Alloys*, Phys. Rev., **125**, no. 2, 541–552, Jan 1962.
4. S. Doniach, *The Kondo lattice and weak antiferromagnetism*, Physica B, **91**, 231–234, Jul 1977.
5. A. C. Hewson, *The Kondo Problem to Heavy Fermions*, Cambridge Studies in Magnetism. Cambridge Universiy Press, Cambridge, 1997.
6. M. A. Ruderman and C. Kittel, *Indirect Exchange Coupling of Nuclear Magnetic Moments by Conduction Electrons*, Phys. Rev., **96**, no. 1, 99, Oct 1954.
7. Tadao Kasuya, *A Theory of Metallic Ferro- and Antiferromagnetism on Zener's Model*, Prog. Theo. Phys., **16**, no. 1, 45–57, 1956.
8. Kei Yosida, *Magnetic Properties of Cu-Mn Alloys*, Phys. Rev., **106**, no. 5, 893–898, Jun 1957.
9. S. Paschen, T. Lühmann, S. Wirth, P. Gegenwart, O. Trovarelli, C. Geibel, F. Steglich, P. Coleman, and Q. Si, *Hall-effect evolution across a heavy-fermion quantum critical point*, Nature, **432**, 881, 2004.
10. M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishnamurthy, *Dynamical cluster approximation: Nonlocal dynamics of correlated electron systems*, Phys. Rev. B, **61**, 12739, 2000.
11. T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, *Quantum cluster theories*, Rev. Mod. Phys., **77**, 1027, 2005.
12. J. E. Hirsch and R. M. Fye, *Monte Carlo Method for Magnetic Impurities in Metals*, Phys. Rev. Lett., **56**, no. 23, 2521–2524, Jun 1986.
13. S. Capponi and F. F. Assaad, *Spin and charge dynamics of the ferromagnetic and antiferromagnetic two-dimensional half-filled Kondo lattice model*, Phs. Rev. B, **63**, 155114, 2001.
14. L. C. Martin and F. F. Assaad, *Evolution of the Fermi Surface across a Magnetic Order-Disorder Transition in the Two-Dimensional Kondo Lattice Model: A Dynamical Cluster Approach*, 2007, arXiv:0711.2235 (submitted to Phys. Rev. Lett.).
15. John A. Hertz, *Quantum critical phenomena*, Phys. Rev. B, **14**, no. 3, 1165–1184, Aug 1976.
16. A. J. Millis, *Effect of a nonzero temperature on quantum critical points in itinerant fermion systems*, Phys. Rev. B, **48**, no. 10, 7183–7196, Sep 1993.
17. Q. Si, S. Rabello, K. Ingersent, and J.L. Smith, *Locally critical quantum phase transitions in strongly correlated metals.*, Nature, **413**, 804, 2001.
18. T. Senthil, Matthias Vojta, and Subir Sachdev, *Weak magnetism and non-Fermi liquids near heavy-fermion critical points*, Physical Review B (Condensed Matter and Materials Physics), **69**, no. 3, 035111, 2004.

Molecular Dynamics Computer Simulation of Multicomponent Liquids: From Effective Potentials to Crystallization from the Melt

Ali Kerrache[1], Antoine Carré [1], Jürgen Horbach[1,2], and Kurt Binder[1]

[1] Institut für Physik, Johannes-Gutenberg-Universität Mainz
Staudinger Weg 7, 55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

[2] Institut für Materialphysik im Weltraum, Deutsches Zentrum für Luft- und Raumfahrt (DLR)
51170 Köln, Germany
E-mail: juergen.horbach@dlr.de

The molecular dynamics (MD) simulation technique is a powerful tool for the investigation of multicomponent liquids and solids. A realistic description of such systems relies on the quality of the effective potential with which the interactions between the atoms are modelled in a MD simulation. We propose a fitting scheme to derive effective potentials from *ab initio* simulations. This scheme is used to parametrize a new potential for silica. In a second case study, MD simulations are used to investigate crystallization in an AlNi alloy, elucidating the crystal growth mechanism on an atomistic scale.

1 Introduction

Molecular dynamics (MD) computer simulations describe the structure and dynamics of condensed matter systems on an atomistic scale[1–4]. In a MD simulation, Newton's equations of motion are solved numerically for an interacting many particle system. As a result, one obtains the trajectories of the particles from which one can calculate, in principle, any static or dynamic quantity of interest. This allows to shed light, e.g., on microscopic transport processes in binary (glassforming) liquids[5].

But how realistic can one model a multicomponent liquid by MD simulation? This depends on the quality of the effective potential that is used to model the interactions between the atoms. To obtain such a potential model for a given system, one usually fits the parameters of some given functional form of the potential to *ab initio* calculations or to experimental data. Very recently, we have proposed a fitting scheme to obtain an effective potential from an *ab initio* MD simulation technique, the so-called Car-Parrinello molecular dynamics (CPMD)[7]. We have used this scheme to parameterize a new pair potential for amorphous silica (SiO_2)[6]. In Sec. 2, we show that our new potential provides an accurate description of the tetrahedral network structure as well as the diffusion dynamics of SiO_2.

With a reliable model potential, one can get insight into microscopic transport mechanisms of real materials. An old problem is to clarify the mechanism of crystal growth from the melt (see, e.g., Ref.[8,9]). In Sec. 3, we consider the crystal growth kinetics of the binary alloy $Al_{50}Ni_{50}$. The experimental melting temperature for this system is at 1920 K where it exhibits a first order phase transition from a liquid to an intermetallic B2 phase. Recent experiments have shown that the crystal growth velocity for this transition is relatively slow[10]. At an undercooling of about 60 K, growth velocities of the order of 0.1 m/s were found. This value is about two orders of magnitude smaller than that found for pure

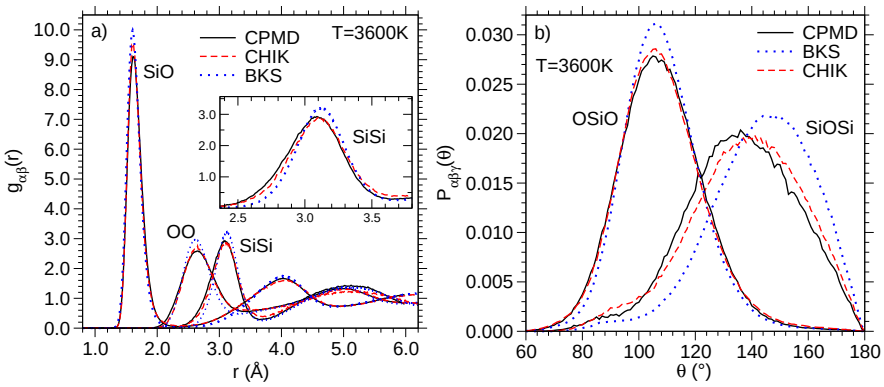


Figure 1. a) Partial pair correlation functions at $T = 3600\,\mathrm{K}$ for the SiSi, SiO, and OO correlations, calculated from CPMD and classical MD using the BKS and the CHIK potentials, as indicated. The inset shows an enlargement of the first peak in $g_{\mathrm{SiSi}}(r)$. b) Angular distribution functions for the OSiO and SiOSi angle at $T = 3600\,\mathrm{K}$, as obtained from CPMD and classical MD with the BKS and the CHIK potential.

metals at comparable undercoolings. The MD simulation sheds light onto the nature of the slow crystal growth in $Al_{50}Ni_{50}$, indicating that diffusion processes in the liquid-crystal interface control the speed of the growth kinetics[11].

2 A New Effective Potential for Silica

In a CPMD simulation, the electronic degrees of freedom are explicitly taken into account on a quantum mechanical level in the framework of density functional theory[12]. Nowadays, systems of 100 to 200 particles can be simulated on a time scale of several tens of picoseconds by CPMD. Apart from studying the structural and dynamic properties of a material, one can also use this technique to derive effective potentials for classical MD simulations, the latter technique providing simulations on much larger scales than CPMD. To obtain an effective potential for a given system, one first has to decide about its functional form. A pair potential model, the so-called BKS potential[13], has been shown to give quite an accurate description of amorphous silica, though it has some deficiencies (e.g., the density of amorphous silica is not well reproduced). In this work, we aimed at developing a pair potential model that is better than the BKS model, in particular with respect to the density. The second step in the development of the potential is to fit the free parameters to the CPMD calculations. This is a non-trivial task since one has to match the CPMD trajectories to a high-dimensional fit parameter space. To this end, we have developed a new iterative fitting scheme that is based on the matching of the partial pair correlation functions from MD runs of the effective potential model with those from CPMD runs. The details of the MD simulations and of the fitting scheme can be found in Refs.[6,14]. As a result, we obtained a new pair potential for silica that we call CHIK potential.

In the following, we compare structural and dynamic quantities, as calculated by CPMD and classical MD, using the BKS and the CHIK model.

Figure 1a shows the partial pair correlation functions[5] at $T = 3600\,\mathrm{K}$, as calculated from CPMD and classical MD using the BKS and the CHIK potential, as indicated. The CHIK model yields good agreement with the CPMD results. The largest differences are

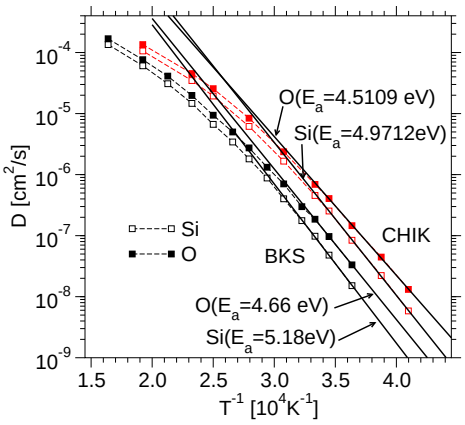


Figure 2. Arrhenius plot of the self-diffusion constants obtained by simulations with the BKS and the CHIK potential, as indicated. The bold solid lines are fits with Arrhenius laws (see text).

found in $g_{\mathrm{SiSi}}(r)$, but also in this case the CHIK potential leads to a better agreement with CPMD than the BKS potential (see the inset in Fig. 1a).

The partial pair correlation functions as obtained from the CPMD runs were used to parameterize the CHIK potential. That these functions are well reproduced by the CHIK potential is thus not that surprising. A less obvious test is provided by considering the distribution functions $P_{\alpha\beta\gamma}(\theta)$ of the bond angles ($\alpha\beta\gamma = \mathrm{OSiO}, \mathrm{SiOSi}$) in comparison to CPMD. As can be inferred from Fig. 1b, the average intra-tetrahedral OSiO angle is around $\theta = 106^\circ$ at 3600 K, both for CPMD and classical MD. However, in the case of the BKS model the function P_{OSiO} is less broad and exhibits a significantly higher amplitude than the CPMD and the CHIK model. Larger differences are seen in the distribution function for the inter-tetrahedral SiOSi angle. While the BKS distribution shows a maximum around $\theta = 147^\circ$, the CHIK model gives a maximum around $\theta = 141^\circ$, in better agreement with the CPMD value around $\theta = 136^\circ$. All the three distribution functions for the SiOSi angle exhibit also a shoulder around $\theta = 90^\circ$ revealing the emergence of edge-sharing tetrahedra at the relatively high temperature $T = 3600\,\mathrm{K}$[15]. In the BKS case, the latter shoulder has a lower amplitude which indicates a smaller number of edge-sharing tetrahedra. Furthermore, as for P_{OSiO} the BKS function for P_{SiOSi} is less broad and its main peak has a higher amplitude. In this sense the BKS model leads to a less disordered structure than the CPMD and the CHIK model.

Detailed comparisons between the CHIK and the BKS model have been also performed with respect to dynamic quantities. Here, we discuss the behaviour of the self-diffusion constants D_α ($\alpha = \mathrm{Si}, \mathrm{O}$). In this case, we take results for the BKS model from a recent simulation study[16]. Whereas the latter results were obtained at the constant density $\rho = 2.37\,\mathrm{g/cm^3}$, the simulation with CHIK potential were done at $\rho = 2.21\,\mathrm{g/cm^3}$ (close to the experimental density of amorphous SiO_2 under ambient conditions which is almost constant over the whole temperature range[17]). The self-diffusion constants were computed from the long-time limit of the mean squared displacements $\langle r_\alpha^2(t)\rangle$ via the Einstein relation $D_\alpha = \lim_{t\to\infty}\langle r_\alpha^2(t)\rangle/6t$[5]. In Fig. 2, the temperature dependence of these transport

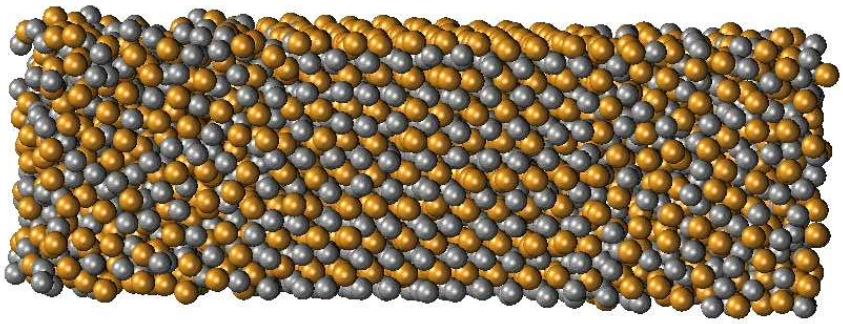


Figure 3. Snapshot of a simulated configuration with two crystal–melt interfaces of the system $Al_{50}Ni_{50}$ at the temperature $T = 1500$ K. Al and Ni atoms are shown as grey and brown spheres, respectively.

coefficients is displayed in an Arrhenius plot. Obviously, the diffusion dynamics, as predicted by the CHIK model, appears to be faster than that of the BKS model. At low temperatures, the self-diffusion coefficients are about a factor of 5 higher than those obtained from the BKS model.

In agreement with various experimental studies[17], at low temperatures the self-diffusion constants can be well-described by an Arrhenius law, $D_\alpha = A_\alpha \exp[-E_\mathrm{a}^\alpha/(k_B T)]$. The activation energies E_a, that we find for the CHIK model, are 4.51 eV for oxygen and 4.97 eV for silicon. These values are very similar to those obtained for the BKS model[16] and they are in good agreement with experimental results ($E_\mathrm{a} = 4.7$ eV for oxygen[18] and $E_\mathrm{a} = 6.0$ eV for silicon[19]). It is remarkable that the temperature dependence of the self-diffusion constants of the CHIK model is very similar to that of the BKS model, although the structure, as obtained from both models, shows significant differences. A dynamic quantity, that is more sensitive to structural differences, is the vibrational density of states. As discussed in Ref.[6], the CHIK model for this quantity yields better agreement with CPMD than the BKS model.

In conclusion, we have developed a new effective potential for silica. The main aim of this work was to develop a fitting scheme for deriving effective (pair) potentials from *ab initio* simulations (CPMD). The new CHIK potential is superior to the BKS model with respect to various static and dynamic properties of amorphous silica. Moreover, as shown in Ref.[6], the CHIK potential has proved to be transferable also to crystalline SiO_2 phases such as α quartz, considering the quantitative agreement between the experimental data for both cell parameters and elastic constants. In the future, our fitting scheme can be used to parameterize potentials for other (amorphous) systems, in particular for mixtures of SiO_2 with other oxides such as Al_2O_3, Na_2O, or MgO.

3 Crystal Growth Kinetics of B2–$Al_{50}Ni_{50}$

To investigate the crystallization of $Al_{50}Ni_{50}$ from the melt, we have done extensive molecular dynamics computer simulations in the NpT ensemble[1–4], i.e., at constant particle number N, constant temperature T and constant pressure ($p_\mathrm{ext} = 0$). The interactions

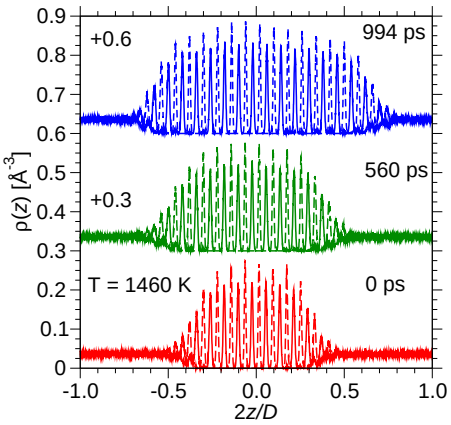


Figure 4. Number density profiles during crystal growth at $T = 1460\,\mathrm{K}$ for Al (dashed lines) and Ni (solid lines). The profiles corresponding to $t = 560\,\mathrm{ps}$ and $t = 994\,\mathrm{ps}$ are shifted with respect to the $t = 0\,\mathrm{ps}$ profiles by 0.3 Å and 0.6 Å, as indicated.

between the atoms were modelled by a potential of the embedded atom (EAM) type, proposed by Mishin *et al.*[20]. Recent studies have shown that this potential gives a realistic description of the diffusion dynamics in Al-Ni melts[21,22] as well as lattice parameters and elastic constants of B2-$Al_{50}Ni_{50}$[20].

At each temperature in the range $1600\,\mathrm{K} \geq T \geq 1200\,\mathrm{K}$, 12 independent samples with solid-liquid interfaces were prepared. To this end, the B2 phase of $Al_{50}Ni_{50}$ was equilibrated at the target temperature for 1 ns. The simulations were done for a system of $N = 3072$ particles ($N_{\mathrm{Al}} = N_{\mathrm{Ni}} = 1536$) in an elongated simulation box of size $L \times L \times D$ (with $D = 3 \times L$), considering the (100) direction of the crystal. Periodic boundary conditions were employed in all three spatial directions. Having relaxed the crystal sample, one third of the particles in the middle of the box were fixed and the rest of the system was melted during 500 ps at $T = 3000\,\mathrm{K}$. Then, the whole system was annealed at the target temperature for another 500 ps, before we started the production runs over 1 ns in the NpT ensemble. A snapshot of the system with two interfaces at $T = 1500\,\mathrm{K}$ is shown in Fig. 3.

The behaviour of samples as the one shown in Fig. 3 depends strongly on the temperature at which they are simulated. While below the melting temperature T_{m}, the crystal will grow, it will melt above T_{m}. From the simulation, the velocity v_{I} with which the liquid-crystal interface moves can be determined. At $T = T_{\mathrm{m}}$, the interface velocity v_{I} vanishes. Thus, by the extrapolation $v_{\mathrm{I}} \to 0$, the melting temperature T_{m} can be estimated. In the following, we show that this procedure yields a rather accurate estimate of T_{m}.

Figure 4 displays the partial number density profiles $\rho(z)$ of Al and Ni at $T = 1460\,\mathrm{K}$ along the z direction, i.e. perpendicular to the solid-liquid interfaces. The lower profiles in Fig. 4 correspond to the starting configuration, while the second and the third ones correspond to $t = 560\,\mathrm{ps}$ and 994 ps. Note that in Fig. 4 the z coordinate is scaled by the factor $2/D$, placing $z = 0$ in the middle of the simulations box. Whereas the crystal structure leads to pronounced peaks in $\rho(z)$, a constant density is observed for the liquid regions

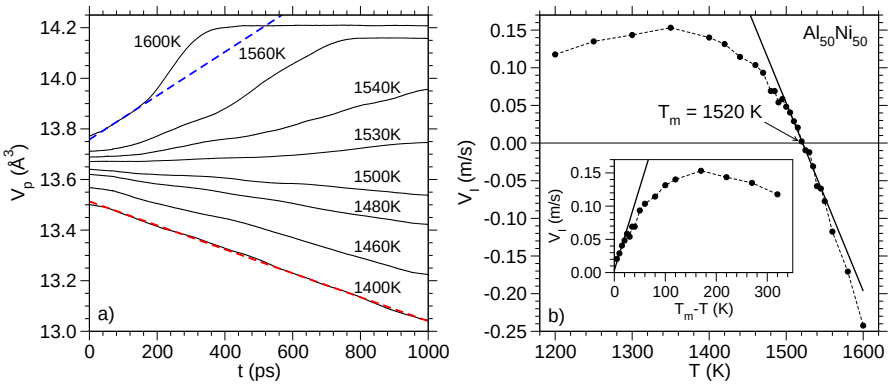


Figure 5. a) Volume per particle, V_p, as a function of time for different temperatures, as indicated. The bold dashed lines are examples of linear fits from which the volume velocity $\dot{V}$ is determined. b) Interface velocity as a function of temperature (filled circles, the dashed line is a guide to the eye). The bold line is a linear fit, $V_\mathrm{I} = k(T_\mathrm{m} - T)$, yielding the kinetic growth coefficient $k = 0.0025\,\mathrm{m/s/K}$ and the melting temperature $T_\mathrm{m} = 1520\,\mathrm{K}$. The inset shows the interface velocity as a function of undercooling $T_\mathrm{m} - T$.

along the z direction, as expected. We can also infer from Fig. 4 that the intermetallic B2 phase (here in (100) orientation) exhibits a pronounced chemical ordering, characterized by the alternate sequence of Al and Ni layers. This indicates that, different from one-component metals, the crystal growth kinetics relies on local rearrangements in the liquid structure. Thus, one may expect that diffusive transport is required to bring the atoms of each species to a suitable site in the B2 crystal. As one can further see in Fig. 4, the crystal is growing at $T = 1460\,\mathrm{K}$. Thus, this temperature is below the melting temperature of our $\mathrm{Al_{50}Ni_{50}}$ model.

Since the density of the crystalline B2 phase is higher than that of the liquid phase, the total volume of the system decreases at temperatures $T < T_\mathrm{m}$ whereas it increases above T_m. Figure 5a shows the time dependence of the volume per particle, V_p, for different temperatures between 1400 K and 1600 K. From this plot, one can infer that the melting temperature is between 1500 K and 1530 K. Also shown in Fig. 5a are examples of linear fits of the form $f(t) = A - \dot{V}_\mathrm{p} t$. Such linear growth laws are expected for steady state growth[23]. We use these fits to determine the change of the volume $\dot{V}$ per unit time. The deviations from the linear behaviour at short times reveal that the growth (or melting) of the crystal is not yet in a steady state regime[23]. At high temperatures, we see a complete melting of the crystal and thus the volume V_p reaches a constant at long times corresponding to the specific volume of the liquid phase. Prior to this, the melting of the crystal is faster than in the linear steady-state regime. In this intermediate regime the crystal has shrunk to such small dimensions that we see essentially the interaction between the two interfaces in the simulation box and thus strong deviations from steady state melting are observed.

From the volume change $\dot{V}_\mathrm{p}$, the velocity v_I, with which the liquid-crystal interfaces move, can be estimated as follows:

$$v_\mathrm{I} = \frac{\dot{V}_\mathrm{p}}{2N_\mathrm{l}(V_\mathrm{c} - V_\mathrm{l})} d \qquad (1)$$

Here, the product $N_\mathrm{l}(V_\mathrm{c} - V_\mathrm{l})$ quantifies the increase of the volume caused by the addition of a crystalline layer (with N_l the average number of particles in a layer, and V_c and V_l the specific volumes of the crystal and the liquid phase, respectively). The length d is the spacing between crystalline layers.

Figure 5b displays the interface velocity v_I as a function of temperature. We see that v_I vanishes around 1520 K and thus this temperature is the estimate for the melting temperature T_m of our simulation model. Note that the experimental value for T_m is around 1920 K and so our simulation underestimates the experimental value by about 20%. Around T_m, the simulation data for v_I can be fitted by the linear law $V_\mathrm{I} = k(T_\mathrm{m} - T)$ where the fit parameter k is the so-called kinetic coefficient. The fit, that is shown in Fig. 5b, yields the value $k = 0.0025$ m/s/K. This value is about two orders of magnitude smaller than the typical values for kinetic coefficients that have been found in simulations of one-component metals[24–27].

The inset in Fig. 5b shows the interface velocity as a function of undercooling $\Delta T = T_\mathrm{m} - T$. We see that v_I increases linearly up to an undercooling of about 30 K. At $\Delta T \approx 180$ K, the interface velocity reaches a maximum value of about 0.15 m/s. Note that our simulation data are in good agreement with recent experimental data on $Al_{50}Ni_{50}$, measured under reduced gravity conditions during a parabolic flight campaign[10].

The very low value of the kinetic coefficient indicates that the crystal growth in $Al_{50}Ni_{50}$ is limited by mass diffusion in the liquid, as proposed by the classical model by Wilson[8] and Frenkel[9]. However, we show in Ref.[11] that, based on sensible assumptions of various parameters, even the Wilson-Frenkel model overestimates the kinetic coefficient for $Al_{50}Ni_{50}$ by at least one order of magnitude. The reason for this can be elucidated by the simulation which reveals that mass diffusion in the liquid-crystal interface controls crystal growth. Since the diffusion in the interface region is much slower than in the bulk liquid, the kinetic coefficient is much lower than expected from the Wilson-Frenkel model.

4 Conclusions

We have presented extensive MD simulation studies that address central issues of materials modelling. The first part was devoted to the development of a new effective potential for silica. We have presented a fitting scheme for deriving effective pair potentials from *ab initio* simulations (CPMD). The method may be used for various molecular systems for which CPMD provides a realistic description. In the second part, the crystallization kinetics of $Al_{50}Ni_{50}$ was investigated. We have shown that the MD simulation gives microscopic insight into the growth kinetics. This is an example that demonstrates the importance of simulation studies, since, at least for metallic alloys, the crystal growth on an atomistic scale is almost not accessible experimentally.

Acknowledgments

A.C. thanks Schott Glas for financial support. The first part of this work is based on collaborative work with Simona Ispas and Walter Kob. For the second part, valuable discussions with Dieter Herlach and Andreas Meyer are gratefully acknowledged. We gratefully

acknowledge financial support within the Priority Program 1120 *Phase Transformations in Multicomponent Melts* of the Deutsche Forschungsgemeinschaft (DFG). Allocation of computer time on the JUMP at the Forschungszentrum Jülich is gratefully acknowledged.

References

1. M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).
2. K. Binder and G. Ciccotti (eds.) *Monte Carlo and Molecular Dynamics of Condensed Matter Systems* (Societá Italiana di Fisica, Bologna, 1996).
3. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 1996).
4. K. Binder, J. Horbach, W. Kob, W. Paul, and F. Varnik, J. Phys.: Condens. Matter **16**, S429 (2004).
5. K. Binder and W. Kob, *Glassy Materials and Disordered Solids: An Introduction to Their Statistical Mechanics* (World Scientific, London, 2005).
6. A. Carré, J. Horbach, S. Ispas, and W. Kob, submitted to Europhys. Lett. (2007).
7. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
8. H.A. Wilson, Phil. Mag. **50**, 238 (1900).
9. J. Frenkel, Phys. Z. Sow. **1**, 498 (1932).
10. S. Reutzel, H. Hartmann, P.K. Galenko, S. Schneider, and D.M. Herlach, Appl. Phys. Lett. **91**, 041913 (2007).
11. A. Kerrache, J. Horbach, and K. Binder, submitted to Europhys. Lett. (2007).
12. W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).
13. B.W.H. van Beest, G.J. Kramer, and R.A. van Santen, Phys. Rev. Lett. **64**, 1955 (1990).
14. A. Carré, L. Berthier, J. Horbach, S. Ispas and W. Kob, J. Chem. Phys. **127**, 114512 (2007).
15. K. Vollmayr, W. Kob and K. Binder, Phys. Rev. B **54**, 15808 (1996).
16. J. Horbach and W. Kob, Phys. Rev. B **60**, 3169 (1999).
17. O.V. Mazurin, M. V. Streltsina. and T.P. Shvaiko-Shvaikovskaya, *Handbook of Glass Data, Part A: Silica Glass and Binary Silicate Glasses* (Elsevier, Amsterdam, 1983).
18. J.C. Mikkelsen, Appl. Phys. Lett. **45**, 1187 (1984).
19. G. Brébec, R. Seguin, C. Sella, J. Bevenot and J.C. Martin, Acta Metall. **28**, 327 (1980).
20. Y. Mishin, M.J. Mehl and D.A. Papaconstantopoulos, Phys. Rev. B **65**, 224114 (2002).
21. S.K. Das, J. Horbach, M.M. Koza., S. Mavila Chathoth and A. Meyer, Appl. Phys. Lett. **86**, 011918 (2005).
22. J. Horbach, S.K. Das, A. Griesche, M.-P. Macht, G. Frohberg and A. Meyer, Phys. Rev. B **75**, 174304 (2007).
23. H.E.A. Huitema, M.J. Vlot and J.P. van der Eerden, J. Chem. Phys. **111**, 4714 (1999).
24. J.J. Hoyt and M. Asta, Phys. Rev. B **65**, 214106 (2002).
25. D.Y. Sun., M. Asta and J.J. Hoyt, Phys. Rev. B **69**, 024108 (2004).
26. Z.G. Xia, D.Y. Sun, M. Asta and J.J. Hoyt, Phys. Rev. B **75**, 012103 (2007).
27. K.A. Jackson, Interface Sc. **10**, 159 (2002).

Ab initio Molecular Dynamics Simulations of the Adsorption of H_2 on Palladium Surfaces

Axel Groß

Institute of Theoretical Chemistry, University of Ulm, 89069 Ulm, Germany
E-mail: axel.gross@uni-ulm.de

The interaction of hydrogen molecules with clean and hydrogen-precovered Pd surfaces was studied using *ab initio* molecular dynamics simulations based on density functional theory. In particular, the dependence of the adsorption probability on the hydrogen coverage and the relaxation of hot hydrogen atoms after dissociation were addressed. The simulations unravel the crucial role of the substrate degrees of freedom in the dissociative adsorption process.

1 Introduction

The interaction of hydrogen with metal surfaces has been one of the benchmark system for the understanding of gas-surface dynamics[1,2]. This is due to the fact that this system is well-suited for both experimental as well as theoretical studies. Furthermore, the hydrogen adsorption on metal surfaces is also relevant from a technological point of view, in particular in the context of hydrogen storage, but also with respect to hydrogenation reaction in heterogeneous catalysis.

Because of the light mass of hydrogen, the dynamics of H_2 adsorption on low-index metal surfaces has mainly been studied using a quantum mechanical treatment on parameterized potential energy surfaces derived from density functional theory (DFT) calculations[3,4]. In these calculations, all six degrees of freedom of the H_2 molecule were explicitly considered, the substrate degrees of freedom, however, were kept frozen since otherwise the calculations would have been computationally much too expensive. This was justified by the large mass mismatch between hydrogen and metal atoms. In addition, comparisons between six-dimensional quantum and classical calculations showed that the role of quantum effects in the hydrogen adsorption dynamics on surfaces is limited[5,6].

If more than six degrees of freedom are involved, quantum dynamical calculations are computationally prohibitive. Furthermore, the parameterization of such high-dimensional potential energy surfaces becomes quite cumbersome. *Ab initio* molecular dynamics (AIMD) simulations of the adsorption dynamics represent an alternative approach. In these simulations, the forces necessary to integrate the equations of motion are determined "on the fly" by first-principles calculations. This does not require any parameterization of the potential energy surfaces. However, up to recently AIMD simulations were restricted to a small number of trajectories because of their high computational effort[7,8].

Previously, we had studied the adsorption and scattering of O_2/Pt(111) using *ab initio* derived tight-binding molecular dynamics (TBMD) simulations[9–11]. This method is about three orders of magnitude faster than AIMD simulations, and it needs relatively little input from *ab initio* total-energy calculations since the tight-binding Hamiltonian already takes the quantum nature of bonding into account[9]. However, the TBMD method still involves a fitting procedure which introduces a source of possible errors; this is true in particular, if more than two different elements are included in the simulations.

Fortunately, due to the increase in computer power and the development of more efficient algorithms it has now become possible to determine a statistically meaningful number of AIMD trajectories[12]. In this contribution, this will be demonstrated using the H_2 adsorption on precovered Pd surfaces and the relaxation of the hydrogen atoms after the dissociation on clean surfaces as examples.

2 Theoretical Methods

Periodic DFT calculations were performed within the generalized gradient approximation (GGA)[13,14] for the exchange-correlation functional using the VASP code[15]. The atomic cores were represented by ultrasoft pseudopotentials[16] allowing a small cutoff energy of 200 eV in the expansion of the one-electron valence states in plane waves. The surface was modelled by slabs of three to five layers for the (100) surface and five layers for the (111) surface. The uppermost layers were allowed to move in the simulations in order to enable recoil and energy transfer processes at the surface.

Adsorption probabilities were determined by averaging over at least 200 trajectories with random initial lateral configurations and molecular orientations for one particular kinetic energy. This leads to a statistical error of $\sqrt{s(1-s)}/\sqrt{200} \leq 0.035$, where s is the adsorption probability. All the substrate atoms were initially at rest, i.e., the initial surface temperature corresponded to 0 K. The MD simulations were performed using the Verlet algorithm with a time step of 1 fs within the microcanonical ensemble. The parallel version of the VASP code was written using a message-passing programming model relying on the MPI library to deal with communications.

3 Hydrogen Dissociation on Hydrogen-Precovered Pd Surfaces

In a recent STM study addressing the formation of ordered hydrogen layers on Pd(111), it was found that hydrogen molecules impinging on an almost complete hydrogen overlayer did not adsorb dissociatively in hydrogen dimer vacancies[17,18]. Instead, aggregates of three or more vacancies were required for the dissociation of H_2. As a consequence, it was speculated that the accepted notion that two empty sites are needed for the dissociative adsorption of diatomic molecules might not be correct[17,19]. However, a subsequent DFT study demonstrated that the dissociative adsorption of H_2 in a hydrogen dimer vacancy is still exothermic[20]. Nevertheless, because of repulsive interactions the presence of the hydrogen overlayer leads to the formation of small energetic barriers. Thus the dissociative adsorption of H_2 on hydrogen-precovered Pd(111) is no longer spontaneous but becomes an activated process by the so-called poisoning effect of the hydrogen overlayer.

Still, the H_2 adsorption process requires the dissipation of more than 1 eV[20]. It has been proposed that the energy transfer to substrate phonons or electron hole-pair excitations could be suppressed due to surface stiffening or the modification of the electronic band structure, respectively, caused by the hydrogen overlayer[21]. The magnitude of these effects, however, could not be assessed so far. Therefore we have addressed the dynamics of the H_2 dissociative adsorption on hydrogen-precovered Pd surfaces by performing AIMD simulations.

In a first step, we used a small initial kinetic energy of 0.02 eV as in the STM experiments. And indeed, we did not find any single adsorption event in a hydrogen dimer

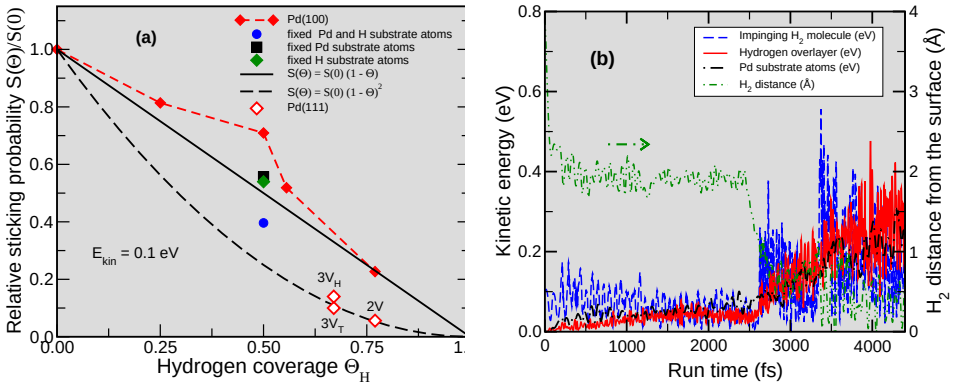


Figure 1. Dissociative adsorption of hydrogen on hydrogen-covered Pd(100) and Pd(111). a) Calculated sticking probability as a function of hydrogen coverage θ_{H} for an initial kinetic energy of 0.1 eV. 2V, $3\mathrm{V_H}$ and $3\mathrm{V_T}$ denote the dimer vacancy and the trimer vacancy centred around a Pd hollow and a Pd top site, respectively. b) Energy redistribution and $\mathrm{H_2}$ centre of mass distance from the surface along a particular trajectory for $\theta_{\mathrm{H}} = 5/9$.

vacancy (2V) on Pd(111) within a (3×3) surface geometry, thus confirming the experiment[17] and providing an explanation for the experimental findings: At the low temperatures of 70 K used in the experiment[17] the impinging $\mathrm{H_2}$ molecules had too little kinetic energy to cross the dissociation barrier.

At higher kinetic energies, however, $\mathrm{H_2}$ molecules can adsorb dissociatively in a dimer vacancy on hydrogen-covered Pd(111). This is shown in Fig. 1a where the sticking probabilities for both Pd(111) and Pd(100) are plotted as a function of the hydrogen coverage normalized to the value at the corresponding clean Pd surfaces for an kinetic energy of 0.1 eV. On Pd(111), a small, but non-vanishing relative adsorption probability of 0.05 is found in the dimer vacancy (2V). At the trimer vacancies on Pd(111) centred either around a hollow site ($3\mathrm{V_H}$) or a Pd top site ($3\mathrm{V_T}$), the adsorption probability is at least twice as large as on the dimer vacancy (2V). Since the area of the trimer vacancy is only larger by 50% compared to the dimer vacancy, this indicates that it is not only the area of the vacancies that determines the adsorption probability.

In Fig. 1a, two curves corresponding to $S(\Theta_{\mathrm{H}}) = S(0)(1 - \Theta_{\mathrm{H}})$ and $S(\Theta_{\mathrm{H}}) = S(0)(1 - \Theta_{\mathrm{H}})^2$ are included which would correspond to the sticking probability if it was determined by pure site-blocking requiring one or two empty sites, respectively. At low and intermediate coverages, the sticking probability of $\mathrm{H_2}$ at Pd(100) is significantly larger than predicted from a simple site-blocking picture, in particular for $\Theta_{\mathrm{H}} = 0.5$. Running additional AIMD trajectories with the substrate kept fixed reveals that it is the energy transfer from the impinging $\mathrm{H_2}$ molecule to the substrate and the rearrangement of the substrate atoms that lead to this enhanced sticking probability compared to pure site blocking[12]. These effects overcompensate the poisoning of the $\mathrm{H_2}$ dissociation caused by the presence of the hydrogen overlayer atoms.

A rather surprising result is, however, that both the motion of the hydrogen atoms and the Pd atoms contribute to the high sticking probability at the precovered surfaces, as the simulations with only fixed Pd atoms and with only fixed hydrogen overlayer atoms show (see the results in Fig. 1a for $\theta_{\mathrm{H}} = 1/2$). Not only energy transfer processes alone but also

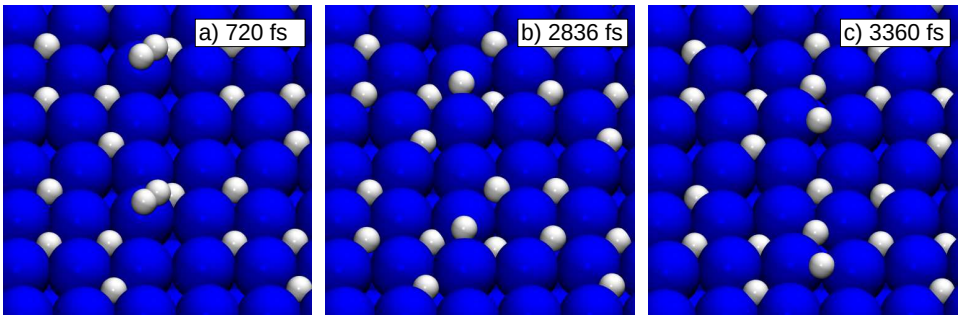


Figure 2. Snapshots of the AIMD trajectory of H_2 impinging on hydrogen-precovered Pd(100) whose energy distribution is plotted in Fig. 1b. The initial hydrogen coverage is $\theta_H = 5/9$ within a (3×3) surface periodicity, i.e., there are four hydrogen vacancies in the surface unit cell.

substrate recoil and relaxation effects in the highly corrugated potential energy surface enhance the sticking probability of H_2 molecules, as a detailed analysis of the adsorption dynamics indicates.

Details of the energy transfer along a particular trajectory for an initial hydrogen coverage of $\theta_H = 5/9$ are presented in Fig. 1b, and three snapshots along this trajectory within the (3×3) periodicity are presented in Fig. 2. This specific example furthermore demonstrates that AIMD simulations do not only yield statistically reliable reaction probabilities, but also allow for valuable microscopic insights into the dissociative adsorption dynamics. At the hydrogen-covered Pd(100) surface, the impinging H_2 molecules do not necessarily either immediately dissociate or scatter back into the gas phase. In fact, there exists a weakly bound molecular precursor state above the top sites which becomes stabilized due to the poisoning effects of the pre-adsorbed hydrogen atoms, very similar to the one already identified at the hydrogen-covered stepped Pd(210) surface[22,23]. This state has not been identified experimentally yet, however, it should be detectable at low surface temperatures by, e.g., isotope exchange experiments since it is bound by 0.1 eV. In the case of the particular trajectory considered in Fig. 1b, after impinging on the surface the H_2 molecule becomes first trapped for more than 2 ps in the molecular precursor state about 2 Å above a Pd atom (Fig. 2a).

At about 2.5 ps, one of the two hydrogen atoms enters the four-fold atomic adsorption site which is associated with an energy gain of about 0.5 eV. This energy is first taken up by the corresponding hydrogen atom but is then rapidly transfered to both the hydrogen overlayer *and* the Pd substrate atoms. The other hydrogen atom remains at a bridge site configuration where is stays for about 0.8 ps (Fig. 2b). The additional bridge-site hydrogen atom can actually move in an exchange mechanism[24] which is what happens after 3.3 ps: the bridge-site hydrogen atom replaces one of the adsorbed hydrogen atoms at an adjacent four-fold hollow sites which is then pushed up to an adjacent empty four-fold hollow site so that eventually all hydrogen atoms end up in the most favourable adsorption sites. However, the replaced hydrogen atom can also end up in another bridge-site configuration which corresponds to an exchange-diffusion of this hydrogen species.

Recently, there has been a renewed interest in the hydrogen absorption in metals[25] in the context of the hydrogen technology. It is important to realize that hydrogen storage still

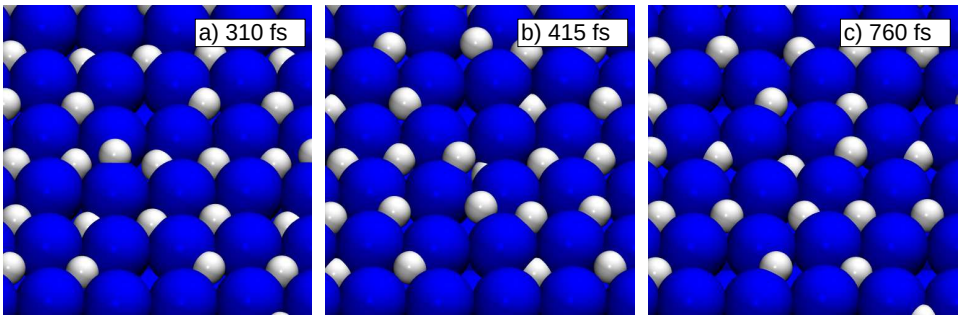


Figure 3. Snapshots of a AIMD trajectory of H_2 impinging on hydrogen-precovered Pd(100) with a kinetic energy of 1.0 eV. Initially, there is a hydrogen dimer vacancy, i.e., the coverage is $\theta_{\mathrm{H}} = 7/9$ within the (3×3) surface periodicity.

remains a problem[26]. Experiments showed that on Pd hydrogen first adsorbs *on* the surface before bulk absorption starts[27]. DFT calculations have confirmed this picture by showing that hydrogen subsurface absorption is energetically less favourable than adsorption on the surface[23]; furthermore, the penetration into the Pd bulk is hindered by a barrier of considerable height[25].

In the AIMD simulations of the hydrogen interaction with hydrogen-precovered Pd surfaces with an initial kinetic energy of 0.1 eV, we did not find any subsurface adsorption events. This is due to the fact that the subsurface penetration corresponds to a rare event because of its activated nature. In order to obtain hydrogen penetration events, we ran trajectories with the unrealistically high initial kinetic energy of 1.0 eV so that the chance is higher that the hydrogen atoms have enough energy to overcome the barrier towards subsurface adsorption. And indeed, thus we observed several hydrogen subsurface adsorption events. Interestingly, most of these events involved a concerted motion of hydrogen atoms which is illustrated in Fig. 3.

Initially, one of the two atoms of the impinging H_2 molecule enters a four-fold hollow site while the other one is again trapped in a bridge configuration (Fig. 3a), as already depicted in Fig. 2b. Note that because of the high initial kinetic energy of 1.0 eV there is a large amount of excess energy available which is mainly transfered to the hydrogen overlayer atoms. After about 400 fs, there is a concerted motion of the hydrogen bridge-site atom and the H atom to its right in Fig. 3b at the adjacent four-fold hollow site. The bridge-site atom replaces the other hydrogen atom, but this time the other atom does not pop up to the surface, but is rather pushed into the subsurface layer. Thus it seems that the subsurface penetration is facilitated when more than one hydrogen atom are involved in this process. The elucidation of the mechanism of the hydrogen subsurface adsorption certainly requires further studies.

After the one hydrogen atom has gone subsurface, the remaining hydrogen atoms on the surface are still energetically hot which is reflected in the large amplitudes of the vibrational motion of the hydrogen atoms with respect to their equilibrium positions that are visible in Figs. 3b and c. Through interatomic collisions, one of the hydrogen overlayer atoms in fact gains enough kinetic energy to cross the diffusion barrier and hops into the neighbouring vacancy site (Fig. 3c).

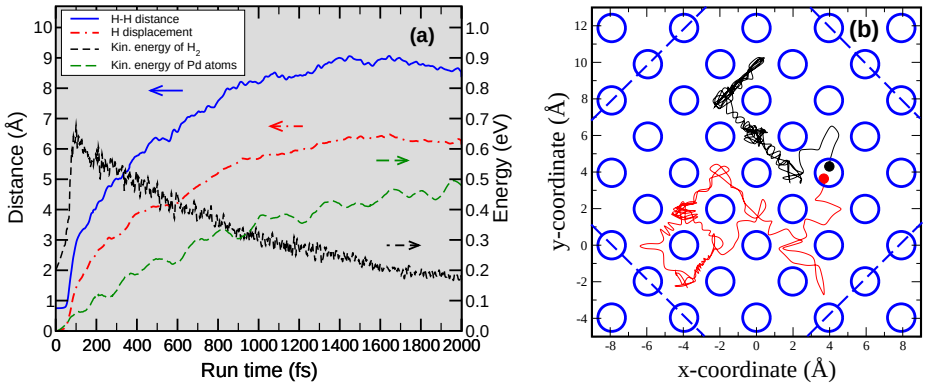


Figure 4. Calculated relaxation of dissociated H_2 molecules on clean Pd(100) within a (6×6) surface unit cell. a) H-H distance, the displacement of the single H atoms from their initial lateral position, the kinetic energy of the H_2 molecules and of the Pd substrate as a function of the run time averaged over 80 AIMD trajectories with an initial H_2 kinetic energy of 200 meV. b) Illustration of a trajectory of the hydrogen atoms after dissociation. The total run time was 2 ps. The surface unit cell of the simulations is indicated by the dashed blue line.

4 Relaxation Dynamics of Dissociated H_2 Molecules

So far we have focused on the dissociation of H_2 on hydrogen-covered Pd surfaces. We were only interested in the question whether H_2 dissociates on Pd or not, but we did not consider the fate of the hydrogen atoms after the dissociation. However, directly after the dissociation when the atoms enter the atomic adsorption wells, they gain a significant amount of energy. For H_2/Pd, this amounts to about 1 eV for the two H atoms together and thus gives rise to the formation of "hot" atoms, i.e., atoms with energies much larger than thermal energies. These atoms can use their kinetic energy in order to travel along the surface. The mean free path of these hot atoms is for example relevant for catalytic reactions on surfaces since it determines whether adjacent reactants can react directly after the dissociative adsorption of one of the species or whether some diffusive motion is required before any further reaction can occur. So far, simulations addressing the relaxation of hot atoms after dissociation have only modelled the motion of single atoms with initial velocities considered to be typical for dissociation fragments directly after the bond-breaking process[28,29]. We have decided to model the full relaxation process including the interaction of the two fragments after the dissociation. We have therefore run AIMD trajectories of the dissociation of H_2 on clean Pd(100). In order to minimize the interaction of the hot hydrogen atoms with their periodic images, we have chosen a large (6×6) surface unit cell. To reduce the computational cost, only three Pd layers were considered but test calculations with five Pd layers yield hardly modified results.

Figure 4a shows the mean H-H distance and the mean displacement of the single H atoms as a function of time averaged over 80 AIMD trajectories whereas in Fig. 4b one specific trajectory is displayed that was run for 2 ps. The trajectory shows that the single hydrogen atoms visit several surface sites. In this particular trajectory, the hydrogen atoms approach each other again after an initial increase in the interatomic distance. At a certain time, they are moving towards adjacent adsorption sites before they separate again. This

shows that the mutual interaction can be important for the hot atom movement.

As for the mean distance between the two H atoms, after about 1.5 ps it reaches a value of about 8.5 Å and does not increase any further. This corresponds to about three Pd lattice units. It should be noted that the root mean square deviation in the H-H distance saturates at about ± 4 Å, however, single trajectories show a maximum H-H distance of more than 20 Å. In Fig. 4a, also the kinetic energies of the hydrogen and the Pd atoms, respectively, are plotted. At about 100 fs, the hydrogen atoms gain on the average about 700 meV when they enter the atomic adsorption well. This energy is then transfered from the hydrogen atoms to the Pd substrate atoms. When the mean H-H distance does not change any more, i.e., after 1.5 ps, the hydrogen atoms have only about 250 meV of kinetic energy left which corresponds rather closely to the diffusion barrier for H atoms on Pd(100).

These results demonstrate that AIMD simulations are well-suited in order unravel details of reaction dynamics at surfaces that are not accessible by experiment.

Acknowledgments

These calculations were made possible through a grant of computer time at the John von Neumann Institute for Computing in Jülich.

References

1. Axel Groß, *Reactions at surfaces studied by ab initio dynamics calculations*, Surf. Sci. Rep., **32**, 291, 1998.
2. G.-J. Kroes, A. Groß, E. J. Baerends, M. Scheffler, and D. A. McCormack, *Quantum theory of dissociative chemisorption on metal surfaces*, Acc. Chem. Res., **35**, 193, 2002.
3. Axel Groß, Steffen Wilke, and Matthias Scheffler, *Six-dimensional quantum dynamics of adsorption and desorption of* H_2 *at Pd(100): steering and steric effects*, Phys. Rev. Lett., **75**, 2718, 1995.
4. G.-J. Kroes, E. J. Baerends, and R. C. Mowrey, *Six-dimensional quantum dynamics of dissociative chemisorption of* ($v = 0$,$j = 0$) H_2 *on Cu(100)*, Phys. Rev. Lett., **78**, 3583, 1997.
5. Axel Groß and Matthias Scheffler, *Ab initio quantum and molecular dynamics of the dissociative adsorption of hydrogen on Pd(100)*, Phys. Rev. B, **57**, 2493, 1998.
6. H. F. Busnengo, E. Pijper, M. F. Somers, G. J. Kroes, A. Salin, R. A. Olsen, D. Lemoine, and W. Dong, *Six-dimensional quantum and classical dynamics study of* H_2 ($v = 0, J = 0$) *scattering from Pd(111)*, Chem. Phys. Lett., **356**, 515, 2002.
7. A. Groß, M. Bockstedte, and M. Scheffler, *Ab initio molecular dynamics study of the desorption of* D_2 *from Si(100)*, Phys. Rev. Lett., **79**, 701, 1997.
8. Lucio Colombi Ciacchi and Mike C. Payne, *"Hot-Atom" O2 Dissociation and Oxide Nucleation on Al(111)*, Phys. Rev. Lett., **92**, 176104, 2004.
9. A. Groß, "Molecular trapping of oxygen at metal surfaces", in: NIC Symposium 2004, D. Wolf, G. Münster, and M. Kremer (Eds.), vol. 20 of *NIC Series*, p. 51, John von Neumann Institute for Computing, Jülich, 2003.
10. A. Groß, A. Eichler, J. Hafner, M. J. Mehl, and D. A. Papaconstantopoulos, *Unified picture of the molecular adsorption process:* O_2*/Pt(111)*, Surf. Sci., **539**, L542, 2003.

11. A. Groß, A. Eichler, J. Hafner, M. J. Mehl, and D. A. Papaconstantopoulos, *Ab initio based tight-binding molecular dynamics simulation of the sticking and scattering of O_2/Pt(111)*, J. Chem. Phys., **124**, 174713, 2006.
12. A. Groß and A. Dianat, *Hydrogen dissociation dynamics on precovered Pd surfaces: Langmuir is still right*, Phys. Rev. Lett., **98**, 206107, 2007.
13. John P. Perdew, J. A. Chevary, S. H. Vosko, Koblar A. Jackson, Mark R. Pederson, D. J. Singh, and Carlos Fiolhais, *Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation*, Phys. Rev. B, **46**, 6671, 1992.
14. John P. Perdew, Kieron Burke, and Matthias Ernzerhof, *Generalized gradient approximation made simple*, Phys. Rev. Lett., **77**, 3865, 1996.
15. G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, Phys. Rev. B, **54**, 11169, 1996.
16. David Vanderbilt, *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism*, Phys. Rev. B, **41**, 7892, 1990.
17. T. Mitsui, M. K. Rose, E. Fomin, D. F. Ogletree, and M. Salmeron, *Dissociative hydrogen adsorption on palladium requires aggregates of three or more vacancies*, Nature, **422**, 705, 2003.
18. T. Mitsui, M. K. Rose, E. Fomin, D. F. Ogletree, and M. Salmeron, *Hydrogen adsorption and diffusion on Pd(111)*, Surf. Sci., **540**, 5, 2003.
19. Miquel Salmeron, *The nature of the catalytic sites for H_2 dissociation*, Topics Catal., **36**, 55, 2005.
20. Nuria Lopez, Zbigniew Lodziana, Francesc Illas, and Miquel Salmeron, *When Langmuir Is Too Simple: H_2 Dissociation on Pd(111) at High Coverage*, Phys. Rev. Lett., **93**, 146103, 2004.
21. S. Holloway, *The active site for dissociative adsorption of H_2: Was Langmuir right?*, Surf. Sci., **540**, 1, 2003.
22. Pia K. Schmidt, Klaus Christmann, Georg Kresse, Jürgen Hafner, Markus Lischka, and Axel Groß, *Coexistence of atomic and molecular chemisorption states: H_2/Pd(210)*, Phys. Rev. Lett., **87**, 096103, 2001.
23. M. Lischka and A. Groß, *Hydrogen adsorption on an open metal surface: H_2/Pd(210)*, Phys. Rev. B, **65**, 075420, 2002.
24. P. J. Feibelman, *Diffusion path for an Al adatom on Al(001)*, Phys. Rev. Lett., **65**, 729, 1990.
25. A. Pundt and R. Kirchheim, *Hydrogen in metals: Microstructural aspects*, Annu. Rev. Mater. Res., **36**, 555–608, 2006.
26. Loius Schlapbach and Andreas Züttel, *Hydrogen-storage materials for mobile applications*, Nature, **415**, 353, 2001.
27. Uwe Muschiol, Pia K. Schmidt, and Klaus Christmann, *Adsorption and absorption of hydrogen on a palladium (210) surface: a combined LEED, TDS, $\Delta\Phi$ and HREELS study*, Surf. Sci., **395**, 182, 1998.
28. C. Engdahl and G. Wahnström, *Transient hyperthermal diffusion following dissociative chemisorption - a molecular-dynamics study*, Surf. Sci., **312**, 429, 1994.
29. N. Pineau, H. F. Busnengo, J. C. Rayez, and A. Salin, *Relaxation of hot atoms following H_2 dissociation on a Pd(111) surface*, J. Chem. Phys., **122**, 214705, 2005.

Boundary Effects in Microfluidic Setups

J. Harting and C. Kunert

Institute for Computational Physics, University of Stuttgart
Pfaffenwaldring 27, 70569 Stuttgart, Germany
E-mail: jens@icp.uni-stuttgart.de

Due to large surface to volume ratios in microfluidic setups, the roughness of channel surfaces must not be neglected since it is not any longer small compared to the length scale of the system. In addition, the wetting properties of the wall have an important influence on the flow. Even though these effects are getting more and more important for industrial and scientific applications, the knowledge about the interplay of surface roughness and hydrophobic fluid-surface interaction is still very limited because these properties cannot be decoupled easily in experiments. We investigate the problem by means of lattice Boltzmann (LB) simulations of rough microchannels with tunable fluid-wall interaction. We introduce an "effective no-slip plane" at an intermediate position between peaks and valleys of the surface and observe how the position of the wall may change due to surface roughness and hydrophobic interactions. We find that the position of the effective wall, in the case of a Gaussian distributed roughness depends linearly on the width of the distribution. Further we are able to show that roughness creates a non-linear effect on the slip length for hydrophobic boundaries.

1 Introduction

The influence of the surface topologies and wetting behaviour of confining geometries in microfluidic systems is of great importance for the understanding of novel techniques using micro- or nanoscale geometries. Such systems allow to handle microliter or nanoliter quantities of liquid for production and analysis processes in the chemical and pharmaceutical industry, for scientific purposes or medical applications. Due to the small length scales in the system, the surface to volume ratio becomes more important. Assuming the surfaces to be perfectly flat and non-interacting is even on molecular scales an invalid assumption which can lead to large errors in experimental measurements. In this report we utilize lattice Boltzmann simulations to investigate the combined influence of roughness and wettability on the fluid flow. This leads to the question which boundary condition has to be applied at a surface in order to treat the surface topology properly. For more than a hundred years the no-slip boundary condition was successfully applied in engineering applications. Nevertheless, Navier[1] introduced a slip boundary condition

$$v(x=0) = \beta \frac{\partial v}{\partial x}$$

saying that the fluid velocity v at the boundary $x = 0$ is proportional to the velocity gradient $\frac{\partial v}{\partial x}$. The constant of proportionality is given by the slip length β. β depends on many parameters like the wettability, the surface roughness or fluid properties like the viscosity or molecular interactions. Therefore, it has to be seen as an empirical length that contains many to some extend unknown interactions. However, for simple liquids the measured slip lengths are commonly of the order of up to some tens of nanometers.

The influence of surface variations on the slip length β has been investigated by numerous authors. On the one hand roughness leads to higher drag forces and thus to no-slip on

macroscopic scales. Richardson showed that even if on a rough surface a full-slip boundary condition is applied, one can determine a flow speed reduction near the boundary resulting in a macroscopic no-slip assumption[2]. This was experimentally demonstrated by McHale and Newton[3]. On the other hand, roughness can cause pockets to be filled with vapour or gas nano bubbles leading to apparent slip[4]. Varnik et al.[5] applied the lattice Boltzmann (LB) method to show that even in small geometries rough channel surfaces can cause flow to become turbulent. Recently, Sbragaglia et al. applied the LB method to simulate fluids in the vicinity of microstructured hydrophobic surfaces[6]. In an approach similar to the one proposed by us, they modelled a liquid-vapour transition at the surface utilising the Shan-Chen multiphase LB model[7]. The authors were able to reproduce the behaviour of the capillary pressure as simulated by Cottin-Bizonne et al. using molecular dynamics (MD) simulations quantitatively[8].

During the last two years, we published a number of papers in which we presented a model to simulate hydrophobic surfaces with a Shan-Chen based fluid-surface interaction and investigated the behaviour of the slip length β[9,10]. We showed that the slip length β is independent of the shear rate, but depends on the pressure and on the concentration of surfactant added. Recently, we presented the idea of an effective wall for rough channel surfaces[11] and investigated the influence of different types of roughness on the position of the effective boundary. Further, we showed how the effective boundary depends on the distribution of the roughness elements and how roughness and hydrophobicity interact with each other[12]. In this report, we revise our previous achievements.

2 Simulation Method

We use a 3D LB model as presented in[13,9] to simulate pressure driven flow between two infinite rough walls that might be wetting or non-wetting. Since the method is well described in the literature we only shortly describe it here.

The lattice Boltzmann equation,

$$\eta_i(\mathbf{x}+\mathbf{c}_i, t+1) - \eta_i(\mathbf{x}, t) = \Omega_i,\ i = 0, 1, \ldots, b\,, \tag{1}$$

with the components $i = 0, 1, \ldots, b$, describes the time evolution of the single-particle distribution $\eta_i(\mathbf{x}, t)$, indicating the amount of quasi particles with velocity $\mathbf{c}_i$, at site $\mathbf{x}$ on a 3D lattice of coordination number $b = 19$, at time-step t.

We choose the Bhatnagar-Gross-Krook (BGK) collision operator

$$\Omega_i = -\tau^{-1}(\eta_i(\mathbf{x}, t) - \eta_i^{\mathrm{eq}}(\mathbf{u}(\mathbf{x}, t), \eta(\mathbf{x}, t))), \tag{2}$$

with mean collision time τ and equilibrium distribution η_i^{eq}[14]. We use the mid-grid bounce back boundary condition and choose $\tau = 1$ in order to recover the no-slip boundary conditions correctly. Interactions between the boundary and the fluid are introduced as mean field body force between nearest neighbours as it is used by Shan and Chen for the interaction between two fluid species[7,9]:

$$\mathbf{F}^{\mathrm{fluid}}(\mathbf{x}, t) \equiv -\psi^{\mathrm{fluid}}(\mathbf{x}, t) g_{\mathrm{fluid,\,wall}} \sum_{\mathbf{x}'} \psi^{\mathrm{wall}}(\mathbf{x}', t)(\mathbf{x}' - \mathbf{x})\,. \tag{3}$$

The interaction constant $g_{\mathrm{fluid,\,wall}}$ is set to 0.08 if not stated otherwise. The wall properties are given by the so-called wall density η_{wall}. This enters directly into the effective mass

$\psi^i = 1 - e^{-\frac{\eta^i}{\eta^0}}$, with the normalized mass $\eta^0 = 1$. With such a model we can simulate slip flow over hydrophobic boundaries with a slip length β of up to 5 in lattice units[9]. It was shown that this slip length is independent of the shear rate, but depends on the interaction parameters and on the pressure.

Here, we model Poiseuille flow between two infinite rough boundaries as shown in Fig. 1. Simulation lattices are 512 lattice units long in flow direction and the planes are

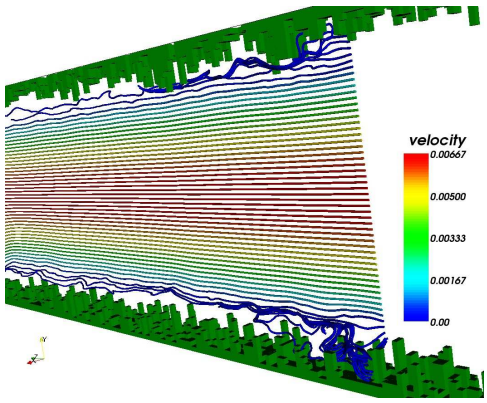


Figure 1. Poiseuille flow in between infinite rough boundaries. The colouring of the streamlines denotes the parabolic velocity profile, while close to the boundary the otherwise laminar streamlines become distorted.

separated by 128 sites between the lowest points of the roughness elements $h_{\min}$. Periodic boundary conditions are imposed in the remaining direction allowing us to keep the resolution as low as 16 lattice units. A pressure gradient is obtained by setting the pressure to fixed values at the in- and outflow boundary. The highest point of one plane gives the height of $h_{\max}$, while the average roughness is found to be R_a (see Fig. 2). In the case of symmetrical distributions $R_a = h_{\max}/2$.

The position of the effective boundary can be found by fitting the parabolic flow profile

$$v_z(x) = \frac{1}{2\mu}\frac{\partial P}{\partial z}\left[d^2 - x^2 - 2d\beta\right] \tag{4}$$

via the distance $2d = 2d_{\text{eff}}$. With β set to 0 we obtain the no-slip case. The viscosity μ and the pressure gradient $\frac{\partial P}{\partial z}$ are given by the simulation. To obtain an average value for d_{eff}, a sufficient number of individual profiles at different positions z are taken into account. The so found d_{eff} gives the position of the effective boundary and the effective height h_{eff} of the rough surface is then defined by $d_{\max} - d_{\text{eff}}$ (see Fig. 2).

3 Flow Along Rough Surfaces

Panzer et al. calculated the slip length β analytically for Poiseuille flow in the case of small cosine-shaped surface variations[15]. It is applicable to two infinite planes separated by a distance $2d$ being much larger than the highest peaks $h_{\max}$. Surface variations are

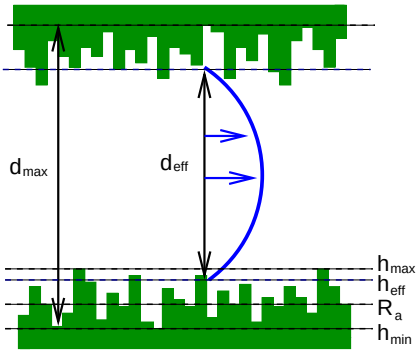


Figure 2. The effective boundary height h_{eff} is found between the deepest valley at h_{min} and the highest peak at h_{max} and corresponds to an effective channel width d_{eff}. For the utilized geometries the average roughness is equal to half the maximum height $R_a = h_{\mathrm{max}}/2$ (from[12]).

determined by peaks of height h_{max}, valleys at h_{min} and given by $h(z) = h_{\mathrm{max}}/2 + h_{\mathrm{max}}/2\cos(qz)$. Here, q is the wave number. Since the surfaces are separated by a large distance, the calculated slip length is equal to the negative effective boundary h_{eff} that is found to be

$$h_{\mathrm{eff}} = -\beta = \frac{h_{\mathrm{max}}}{2}\left(1 + k\frac{1 - \frac{1}{4}k^2 + \frac{19}{64}k^4 + \mathcal{O}(k^6)}{1 + k^2(1 - \frac{1}{2}k^2) + \mathcal{O}(k^6)}\right). \tag{5}$$

The first and k independent term shows the linear behaviour of the effective height h_{eff} on the average roughness $R_a = h_{\mathrm{max}}/2$. Higher order terms cannot easily be calculated analytically and are neglected. Thus, Eq. 5 is valid only for $k = qh_{\mathrm{max}}/2 \ll 1$. However, for realistic surfaces, k can become substantially larger than 1 causing the theoretical approach to fail. Here, only numerical simulations can be applied to describe arbitrary boundaries. In Fig. 3 the normalized effective height h_{eff}/R_a obtained from our simulations is plotted versus k for cosine shaped surfaces with $h_{\mathrm{max}}/2 = k = 1, \frac{1}{2}, \frac{1}{3}$ (symbols). The line is given by the analytical solution of Eq. 5. For $k < 1$ the simulated data agrees within 2.5% with Panzer's prediction. However, for $k = 1$ a substantial deviation between numerical and analytical solutions can be observed because Eq. 5 is valid for small k only. In the case of large $k > 1$, the theory is not able to correctly reproduce the increase of β with increasing h_{max} anymore. Instead, $2\beta/h_{\mathrm{max}}$ becomes smaller again due to missing higher order contributions in Eq. 5. Our simulations do not suffer from such limitations allowing us to study arbitrarily complex surface geometries[11].

We showed that the position of the effective boundary height is depending on the shape of the roughness elements, i.e., for strong surface distortions it is between 1.69 and 1.90 times the average height of the roughness $R_a = h_{\mathrm{max}}/2$[11]. By adding an additional distance between roughness elements, h_{eff} decreases slowly, so that the maximum height is still the leading parameter. We are also able to simulate flow over surfaces generated from AFM data of gold coated glass used in microflow experiments by O.I. Vinogradova and G.E. Yakubov[16]. We find that the height distribution of such a surface is Gaussian and that a randomly arranged surface with a similar distribution gives the same result for the position of the effective boundary although in this case the heights are not correlated. We

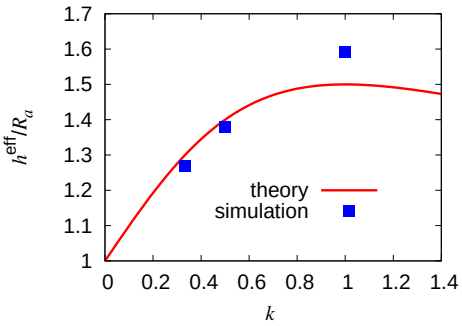


Figure 3. Effective height h_{eff} normalized by the average roughness R_a versus $k = h_{\max}/2q$ for a cosine geometry. Symbols denote numerical data and the line is given by Eq. 5. For $k > 1$ the theory fails simulations are still valid in this regime (from[12]).

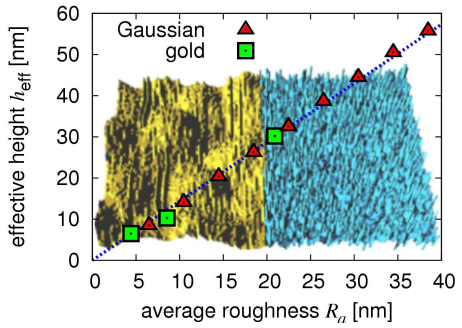


Figure 4. Simulated h_{eff} versus R_a for gold coated glass and a randomly generated surface with Gaussian distributed heights. The background image shows the gold surface (left) and the artificially generated structure (right)[11].

can set the width of the distribution σ and the average height R_a. By scaling σ with R_a we obtain geometrically similar geometries. This similarity is important because the effective height h_{eff} scales with the average roughness in the case of geometrical similarity[11] (see Fig. 4). As an extension of our previous work, we investigate Gaussian distributed heights with different widths σ. In Fig. 5 the effective height h_{eff} is plotted over the average height R_a for $0.054 < \sigma/R_a < 0.135$. The height of the effective wall depends linearly on σ in the observed range as can be seen in the inset[12]. The effective height h_{eff} ranges from $1.15R_a$ to $1.45R_a$. These values are lower than the effective heights for an equally distributed roughness ($1.84R_a$).

4 Wettability and Roughness

We also investigate how roughness and the surface wettability act together by performing simulations with rough channels to which we assign a fluid-wall interaction as given in the

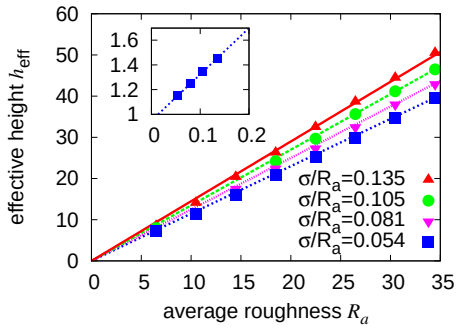


Figure 5. Effective height h_{eff} over average roughness R_a for Gaussian distributed height elements with different width of the distribution σ. Symbols are the simulation results, lines are a linear fit to the data. The inset shows the linear dependence of the effective height on σ[12].

introduction (Eq. 3, η_{wall}= 0.5, 1, and 5). For perfectly smooth surfaces we determine β to be 0.65, 1.13, and 1.3. Fig. 6 depicts the effective height of rough hydrophobic walls versus R_a. For $R_a > 4$ we find a linear dependence between R_a and h_{eff}. The slope for different η_{wall} varies because the fluid-surface interaction does not cause a simple offset on the effective height h_{eff}. Instead, non-linear effects are playing a role.

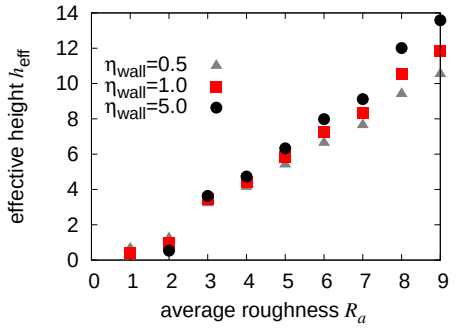


Figure 6. Effective height h_{eff} versus average roughness R_a with different fluid-wall interaction constant η_{wall}. The position of the effective height h_{eff} spreads wider for higher R_a, because larger roughness increases the fluid-wall interaction[12].

To decouple the effects of roughness and wettability we determine the slip length by setting the effective distance d_{eff} in equation (4) to the effective distance for a rough no-slip wall. We then fit the corresponding velocity profile via β. In Fig. 7 we can see that the slip length β for the strong fluid-wall interaction ($\eta^{\mathrm{wall}} = 5$) first decreases with the average roughness and then rises. For a lower interaction, the slip length is constantly growing and leads to an increase of the slip length for weak fluid wall interaction ($\eta_{\mathrm{wall}} = 0.5$) by a factor of more than three. There are two counteracting effects in this system and their interplay can explain the observed behaviour. The decrease of the slip length β is due to an increased friction near the boundary at moderate roughness. The increase has its

reason in the reduced pressure near the hydrophobic rough surface, so that the fluid "feels" a smoothed effective surface. For a more detailed study on superhydrophobic surfaces, the strong surface variation as well as the liquid-gas transitions have to be taken into account. This is ongoing work and will be reported on in the future.

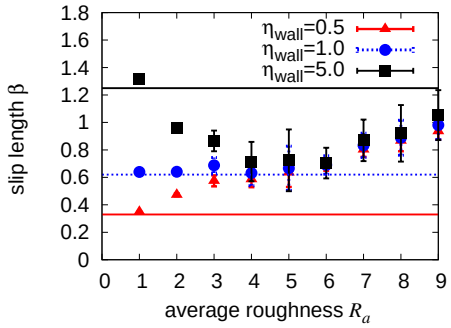


Figure 7. Slip length β over average roughness R_a for equally distributed height elements with different fluid-wall interaction $\eta_{\mathrm{wall}} = 0.5,\ 1.0,\ 5.0$. The position of the effective height h_{eff} is chosen as the value for a non-interacting wall. The lines show the slip lengths for smooth boundaries. Error bars show the standard deviation of results from four different random surfaces[12].

5 Conclusion

In this report we summarized our work on fluid flow along rough and hydrophobic surfaces which has been performed during the last two years. We demonstrated that there is a linear dependence of the effective height on the average roughness and that the average height scales linearly with the width of the distribution of heights σ. We successfully applied our simulations to experimental data and showed that neglecting roughness can lead to substantial errors in experimental measurements. Currently, we investigate the interplay between roughness and hydrophobic fluid-wall interactions and presented preliminary results. They show that there exist non-linear interactions between roughness and hydrophobicity leading to an increase of the slip length and eventually to superhydrophobic effects.

Acknowledgments

We thank the John von Neumann Institute for Computing, Jülich for providing the computing time and technical support for the presented work. This work was financed within the DFG priority program "nano- and microfluidics" and by the "Landesstiftung Baden-Württemberg".

References

1. C.L.M.H. Navier, *Mémoire sur les lois du mouvement de fluids*, Mem. Acad. Sci. Ins. Fr., **6**, 389, 1823.

2. S. Richardson, *On the no-slip boundary condition*, J. Fluid Mech., **59**, 707, 1973.
3. G. McHale and M.I. Newton, *Surface roughness and interfacial slip boundary condition for quarzcrystal microbalances*, J. Appl. Phys., **95**, 373, 2004.
4. P. Joseph, C. Cottin-Bizonne, J. M. Benoi, C. Ybert, C. Journet, P. Tabeling, and L. Bocquet, *Slippage of water past superhydrophobic carbon nanotube forest in microchannels*, Phys. Rev. Lett., **97**, 156104, 2006.
5. F. Varnik, D. Dorner, and D. Raabe, *Roughness-induced flow instability: A lattice Boltzmann study*, J. Fluid Mech., **573**, 191, 2006.
6. M. Sbragaglia, R. Benzi, L. Biferale, S. Succi, and F. Toschi, *Surface roughness-hydrophobicity coupling in micro and nanochannel flows*, Phys. Rev. Lett., **97**, 204503, 2006.
7. X. Shan and H. Chen, *Lattice boltzmann model for simulating flows with multiple phases and components*, Phys. Rev. E, **47**, 1815, 1993.
8. C. Cottin-Bizone, C. Barentin, E. Charlaix, L. Bocquet, and J.L. Barrat, *Dynamics of simple liquids at heterogeneous surfaces: molecular dynamics simulations and hydrodynamic description*, Eur. Phys. J. E, **15**, 427, 2004.
9. J. Harting, C. Kunert, and H.J. Herrmann, *Lattice Boltzmann simulations of apparent slip in hydrophobic microchannels*, Europhys. Lett., pp. 328–334, 2006.
10. C. Kunert and J. Harting, *On the effect of surfactant adsorption and viscosity change on apparent slip in hydrophobic microchannels*, Progress in CFD, **in press**, 2007.
11. C. Kunert and J. Harting, *Roughness induced apparent boundary slip in microchannel flows*, Phys. Rev. Lett., **99**, 176001, 2007.
12. C. Kunert and J. Harting, *Simulation of fluid flow in hydrophobic rough micro channels*, Submitted for publication, **arXiv:0709.3966**, 2007.
13. J. Harting, M. Harvey, J. Chin, M. Venturoli, and P. V. Coveney, *Large-scale lattice Boltzmann simulations of complex fluids: advances through the advent of computational grids*, Phil. Trans. R. Soc. Lond. A, **363**, 1895–1915, 2005.
14. S. Succi, *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*, Oxford science publications, 2001.
15. P. Panzer, M. Liu, and D. Einzel, *The effects of boundary curvature on hydrodynamic fluid flow:Calculation of slip lengths*, Int. J. mod. Phys. B, **6**, 3251, 1992.
16. O.I. Vinogradova and G.E. Yakubov, *Surface roughness and hydrodynamic boundary conditions*, Phys. Rev. E, **73**, 045302(R), 2006.

Free-Energy Barriers of Spin Glasses

Elmar Bittner, Andreas Nußbaumer, and Wolfhard Janke

Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig
Postfach 100 920, 04009 Leipzig, Germany
E-mail: wolfhard.janke@itp.uni-leipzig.de

The Ising spin glass in the Sherrington-Kirkpatrick (SK) mean field and the three-dimensional Edwards-Anderson (EA) nearest-neighbour formulations are investigated by means of Monte Carlo simulations. To this end, we employ a combination of the multioverlap algorithm with parallel tempering methods. In this report we focus on the finite-size scaling behaviour of the free-energy barriers which are visible in the probability density of the Parisi overlap parameter. Assuming that the mean barrier height diverges with the number of spins N as N^{α}, our data for the SK model show good agreement with the theoretical prediction $\alpha = 1/3$. We compare the scaling behaviour to the data from the EA model.

1 Introduction

A major open problem in statistical physics is the nature of the "glassy" low-temperature phase of finite-dimensional spin-glass systems. It is still unresolved whether the replica symmetry-breaking theory or the phenomenological droplet picture yields the correct description (for reviews, see Refs. 1–4). Even at the mean-field level, only very recently a mathematical proof[5] of Parisi's replica solution[6] for the Sherrington-Kirkpatrick (SK) model[7] was given.

In the thermodynamic limit the frozen phase of the mean field spin glass shows many stable and metastable states. Such a feature is the consequence of the disorder and the frustration characterising spin glasses in general, leading to a rugged free-energy landscape with probable regions (low free energy) separated by rare-event states (high free energy). But also for finite systems the free-energy landscape shows an intricate, corrugated structure. Therefore, it is hard to measure the free-energy barriers by means of conventional Monte Carlo simulations directly. The aim of this project is to study the free-energy barriers of the SK mean field spin-glass model and the three-dimensional Edwards-Anderson (EA) nearest-neighbour model[8] using a combination of the multioverlap Monte Carlo algorithm[9] with parallel tempering methods[10]. By using this combined algorithm we are able to perform simulations at much lower temperatures for the EA model than in previous studies[11]. This is necessary, because for temperatures close to the spin-glass transition significant deviations from the theoretical mean-field prediction were found in both the three- and four-dimensional EA model. Since one possible explanation for these deviations are strong finite-size effects close to the spin-glass transition, by measuring at lower temperatures these effects should become less pronounced.

2 Model and Simulation Techniques

The Hamiltonian of the Sherrington-Kirkpatrick mean-field model reads

$$H_{\mathrm{SK}} = -\sum_{i<j} J_{ij} s_i s_j \ , \tag{1}$$

where $s_i = \pm 1$, $i = 1,\ldots,N$, with N denoting the number of spins. The exchange coupling constants J_{ij} are quenched, independent random variables with a Gaussian distribution of zero mean and variance N^{-1}. The critical temperature of the infinite system is $T_c = 1$. The SK model gives a mean-field formulation of spin glasses, while the Edwards-Anderson model describes a finite-dimensional spin glass. Here, the sum in the Hamiltonian runs only over nearest-neighbour pairs of Ising spins,

$$H_{\mathrm{EA}} = -\sum_{\langle ij \rangle} J_{ij} s_i s_j \ , \tag{2}$$

where we consider a simple cubic lattice with $N = L^3$ spins and periodic boundary conditions. In the EA model we draw the exchange coupling constants from a symmetric bimodal distribution such that $J_{ij} = \pm 1$, with equal probability. The spin-glass transition temperature in the three-dimensional EA model was found to be $T_c \approx 1.15$.[12]

The fact that there is no explicit order parameter which allows one to exhibit the free-energy barriers led us to use the Parisi overlap parameter[6],

$$q = \frac{1}{N} \sum_{i=1}^{N} s_i^{(1)} s_i^{(2)} \ , \tag{3}$$

where the spin superscripts label two independent (real) replicas for the same realization of randomly chosen exchange coupling constants $\mathcal{J} = \{J_{ij}\}$. For given $\mathcal{J}$ the probability density of q is denoted by $P_{\mathcal{J}}(q)$, and the function $P(q)$ is obtained as

$$P(q) = [P_{\mathcal{J}}(q)]_{\mathrm{av}} = \frac{1}{\#J} \sum_{\mathcal{J}} P_{\mathcal{J}}(q) \ , \tag{4}$$

where $[\ldots]_{\mathrm{av}}$ symbolises the quenched average and $\#J$ is the number of realizations considered. For a given realization of $\mathcal{J}$ the nontrivial (i.e., away from $q = \pm 1$) minima are related to the free-energy barriers of this disordered system. We are, therefore, interested in the whole range of the probability density $P_{\mathcal{J}}(q)$. Conventional, canonical Monte Carlo simulations are not suited for such systems because the likelihood to generate the corresponding rare-event configurations in the Gibbs canonical ensemble is very small. This problem can be overcome by non-Boltzmann sampling[13,14] with the multi-overlap weight[9]

$$w_{\mathcal{J}}(q) = \exp\left[\beta \sum_{ij} J_{ij} \left(s_i^{(1)} s_j^{(1)} + s_i^{(2)} s_j^{(2)} \right) + S_{\mathcal{J}}(q) \right] \ , \tag{5}$$

where the sum runs over all pairs of spins for the SK model and only over nearest-neighbour pairs for the EA model. The two replicas are coupled by $S_{\mathcal{J}}(q)$ in such a way that a broad multi-overlap histogram $P_{\mathcal{J}}^{\mathrm{muq}}(q)$ over the entire accessible range $-1 \le q \le 1$ is obtained. When simulating with the multi-overlap weight, canonical expectation values of any quantity $\mathcal{O}$ can be reconstructed by reweighting,

$$\langle \mathcal{O} \rangle_{\mathcal{J}}^{\mathrm{can}} = \langle \mathcal{O} \exp(-S_{\mathcal{J}}) \rangle_{\mathcal{J}} / \langle \exp(-S_{\mathcal{J}}) \rangle_{\mathcal{J}} \ . \tag{6}$$

Ideally the weight function $W_{\mathcal{J}} \equiv \exp(S_{\mathcal{J}})$ should satisfy

$$P_{\mathcal{J}}^{\mathrm{muq}}(q) = P_{\mathcal{J}}^{\mathrm{can}} W_{\mathcal{J}} = \mathrm{const.} \ , \tag{7}$$

i.e., should give rise to a completely flat multi-overlap probability density $P_{\mathcal{J}}^{\mathrm{muq}}(q)$. Of course, $P_{\mathcal{J}}^{\mathrm{can}}(q)$ is a priori unknown and one has to proceed by iteration. An efficient way to construct the weight function $W_{\mathcal{J}}$ is to use an accumulative recursion, in which the new weight factor is computed from all available data accumulated so far[15,16]. The multi-overlap algorithm combined with this recursion allows an almost automatic simulation.

The efficiency of the multi-overlap algorithm decreases with lowering the temperature, and since we are mainly interested in the low-temperature behaviour of the spin-glass models we had to seek for suitable algorithmic improvements. As a result of this investigation, we developed a combination of the multi-overlap algorithm[9] with the parallel tempering (PT) update scheme[10], where configurations simulated at different temperatures are tried to be exchanged via a Monte Carlo process that typically follows the Metropolis acceptance rules.[17] This renders the combined algorithm particularly suited for a parallel computer such as JUMP at NIC Jülich. For the PT procedure we used a set of N_T temperature values in the range $T_{\min} < T_c < T_{\max}$. Once for each temperature the entire range of q was covered, the accumulative recursion for the weight functions was stopped. Due to large differences in the free-energy landscape for different disorder realizations $\mathcal{J}$, the number of recursion steps varied for different $\mathcal{J}$. After the weight functions were constructed, they were kept fixed and we took our measurements. Thereby, we recorded time series of the overlap parameter q and the canonical $P_{\mathcal{J}}(q)$ distribution for all temperature values, for an EA example see Fig. 1.

3 Analysis Method

To analyse the low temperature behaviour we computed the barrier autocorrelation time τ_B^q for each of these samples $\mathcal{J}$ by employing the same method as Berg *et al.*[11] used for the EA Ising spin-glass model. For clarity, we recall the basic idea here. The free-energy barrier F_B^q for a given $P_{\mathcal{J}}(q)$ is defined through the autocorrelation time of a one-dimensional Markov process which has the canonical $P_{\mathcal{J}}(q)$ distribution as equilibrium state. The transition probabilities $T_{i,j}$ are given by

$$T = \begin{bmatrix} 1-w_{2,1} & w_{1,2} & 0 & \dots \\ w_{2,1} & 1-w_{1,2}-w_{3,2} & w_{2,3} & \dots \\ 0 & w_{3,2} & 1-w_{2,3}-w_{4,3} & \dots \\ 0 & 0 & w_{4,3} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} , \tag{8}$$

where $w_{i,j}$ ($i \neq j$) is a probability à la Metropolis to jump from state $q = q_j$ to $q = q_i$ ($q_i = i/N$, $i \in [-N, -N+2, \dots, +N]$),

$$w_{i,j} = \frac{1}{2} \min\Big(1, \frac{P_{\mathcal{J}}(q_i)}{P_{\mathcal{J}}(q_j)}\Big) . \tag{9}$$

The transition matrix T fulfills the detailed balance condition (with $P_{\mathcal{J}}$), and as a consequence it has only real eigenvalues. The largest eigenvalue (equal to one) is non-degenerate, and the second largest eigenvalue λ_1 determines the autocorrelation time of the Markov chain,

$$\tau_B^q = -\frac{1}{N \log(\lambda_1)} . \tag{10}$$

The associated free-energy barrier for realization $\mathcal{J}$ is defined as

$$F_B^q = \ln(\tau_B^q) \; . \tag{11}$$

Note that the definition of the autocorrelation time (10) takes only barriers in q into account, but not other barriers which may well exist in the multidimensional configuration space.

4 Numerical Details and Results

Let us start with the SK model where we studied systems with $N = 32, 64, 128, 256, 512,$ and 1024 spins and used a set of 32 temperature values in the range $T = 1/3 - 1.6$ for all of our systems apart from the largest, where we used 64 temperature values for the same temperature interval. We took about $100\,000$ measurements, with five sweeps between the measurements. A sweep consisted of N spin flips with the multioverlap algorithm and one parallel tempering update. To average over the disorder we used 1000 realizations of the disorder for $N \leq 512$ and 100 for $N = 1024$.

For the EA model we studied systems with $N = 4^3, 6^3, 8^3, 10^3,$ and 12^3 spins within a temperature range of $T = 0.5$ – 1.5. Up to $N = 10^3$ we used 11 equally spaced temperature values and for the largest system the spacing below $T = 1$ was halved, leading to 16 replicas. Due to the larger autocorrelation times for the EA model we took at least 10^7 measurements and for the disorder average we used more than 1000 realizations.

For each temperature value we performed least-squares fits of the finite-size scaling (FSS) ansatz $F_B^q = cN^\alpha$ which corresponds to the exponential FSS behaviour

$$\tau_B^q \propto e^{cN^\alpha} \; . \tag{12}$$

The results for the SK model[18] depicted in Fig. 2 are consistent with previous results in the literature[19–26] using analytical and different numerical methods. The horizontal line in Fig. 2 indicates the theoretical prediction $\alpha = 1/3$ of Ref. 24. The figure shows fits with different lower bounds $N_{\min}$ of the fit range, while the upper bound was always our largest system $N = 1024$. From these fits we observe a strong finite-size effect for $T \to T_c = 1$. At lower temperatures we find a linearly increasing deviation from the theoretical value. This is presumably also a finite-size effect, because the slope of the deviation becomes flatter when increasing the lower bound of the fit range and there is no physical reason for a change of behaviour of the barrier autocorrelation time in the glassy phase. Using ansatz (12) for the EA model, the value of the exponent α varies from 0.49 to 0.46 in the interval $T_c \geq T \geq 0.8$, but the quality of the fits is unacceptably low. Only for the smallest temperatures the goodness-of-fit parameter Q is significantly larger then zero, c.f. Fig. 2. We therefore performed power-law fits as in Ref. 11,

$$\tau_B^q = cN^\alpha, \tag{13}$$

which corresponds to a fit of the form $F_B^q = \log(c) + \alpha \log(N)$. The Q-values for these fits are much closer to unity, see Fig. 2. Our data favour them strongly over the exponential finite-size scaling behaviour (12), which confirms previous results[11] for $T = 1$ and extends them to considerably lower temperatures.

One possible explanation for this deviation from the theoretical value is the lack of self-averaging of the finite volume Parisi overlap parameter distribution $P_{\mathcal{J}}$ in the SK model.[27] This has been confirmed numerically for the SK model[18] as well as for the EA model[11].

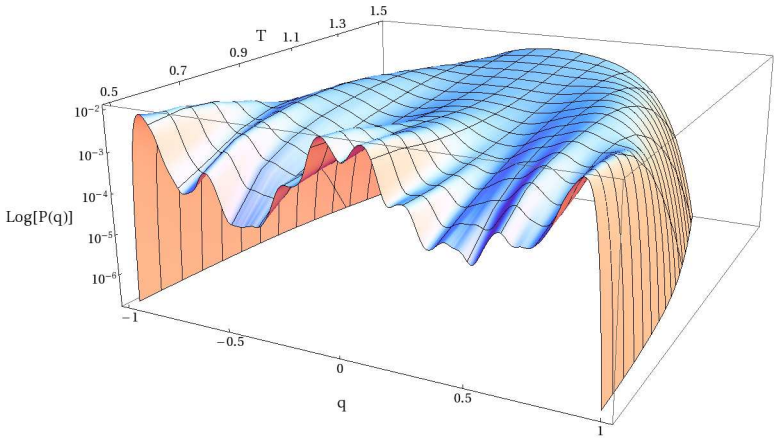


Figure 1. EA model: The logarithm of the canonical $P(q)$ distribution for a 8^3 lattice as a function of temperature for a typical disorder realisation.

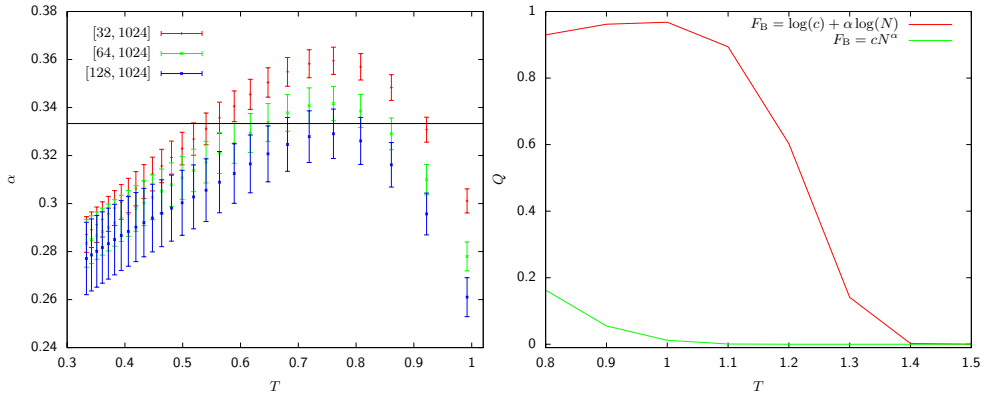


Figure 2. SK model (left plot): Dependence of the exponent α on the lower bound of the fit range $[N_{\rm min}, 1024]$ as a function of temperature. The horizontal line indicates the theoretical value $\alpha = 1/3$. EA model (right plot): Goodness-of-fit parameter Q as a function of temperature for different types of fits.

We already mentioned that the distribution of the free-energy barriers becomes broader for low temperatures. In recent work Dayal *et al.* have found that the tunnelling times of their flat-histogram sampling simulations of the 2D $\pm J$ Ising spin glass are distributed according to the Fréchet extreme-value distribution for fat-tailed distributions.[28] In general, extreme-value statistics can be classified into different universality classes[29,30], depending on whether the tails of the original distribution are fat tailed (algebraic), exponential, or thin tailed (decaying faster then exponential). Assuming that the tunnelling times respectively free-energy barriers are distributed according to an extreme-value distribution, we use the

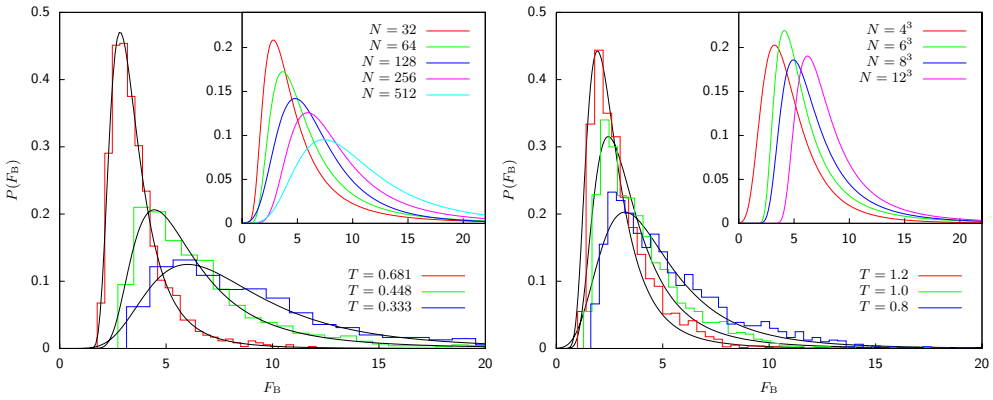


Figure 3. SK model (left plot): Distribution of free-energy barriers F_B^q for $N = 256$ at different temperatures. The inset shows the distribution for $T = 1/3$ for different numbers of spins. EA model (right plot): The same graph as for the SK model for $N = 4^3$ at different temperatures and the inset shows the distribution for $T = 0.8$ for different lattice sizes.

integrated probability density of the generalized extreme-value distribution (GEV),

$$F_{\xi;\,\mu;\,\sigma}(x) = \exp\left[-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right] \tag{14}$$

for $1 + \xi(x - \mu)/\sigma > 0$, to fit our data. We find that the free-energy barriers show fat tails for $T < T_c$ with shape parameter $\xi > 0$, i.e., a Fréchet distribution. In Fig. 3 we plot the resulting distribution for the SK model for different temperatures below the spin-glass transition and find that the tails become fatter and fatter as the temperature goes to zero. The histograms for low temperatures show deviations from the Fréchet distribution for small values of F_B^q, so a much larger number of disorder realizations would be needed to determine both tails of the distribution properly. We determined the parameters σ, μ and ξ for different temperatures and found that σ grows linearly and μ logarithmically with inverse temperature $1/T$, whereas ξ stays more or less constant at $\xi \approx 0.33$. As an example we show in Fig. 4 the results for $N = 512$. If we keep the temperature fixed and look at the size dependence of the distribution, we find that for a larger number of spins the distribution becomes broader, c.f. the inset of Fig. 3. To quantify this behaviour we use the scaling relations $\sigma \propto N^{\alpha(\sigma)}$ and $\mu \propto N^{\alpha(\mu)}$, which lead to $\alpha(\sigma) \approx 0.25$ and $\alpha(\mu) \approx 0.31$ for our lowest temperatures, see the inset of Fig. 4. We find a temperature dependence of the exponents $\alpha(\sigma)$ and $\alpha(\mu)$ with negative and positive slope for increasing T, respectively. For the EA model we also find fat-tailed distributions, but the broadening of the distribution with increasing number of spins is much weaker than for the SK model, see Fig. 3.

5 Conclusion

We found that the free-energy barriers of the SK model are non-self-averaging and distributed according to the Fréchet extreme-value distribution. These particular features were

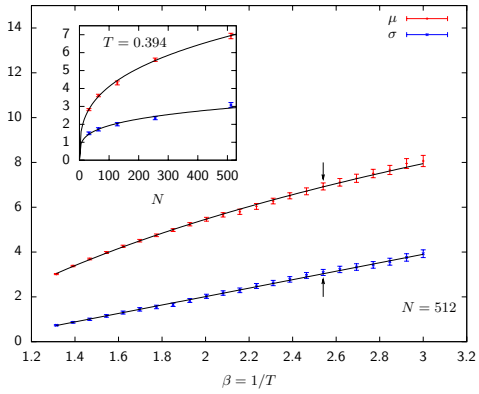


Figure 4. SK model: Temperature dependence of the parameters σ and μ of the Fréchet distribution for $N = 512$. The inset shows the size dependence of σ and μ for $T = 0.394$, indicated by the arrows.

also found for the EA nearest-neighbour model and such similarities support the position that the Parisi replica symmetry breaking solution of the SK model is the limit of the short-range model on a lattice in dimension d when $d \to \infty$, with a proper rescaling of the strength of the Hamiltonian. On the other hand, we also found that the free-energy barriers diverge with the theoretically predicted value $\alpha = 1/3$, which is in contrast to our results for the EA model in three dimensions and previous findings for the three- and four-dimensional EA model[11].

Acknowledgments

We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) under Grants No. JA 483/22-1 and No. JA 483/23-1, the EU RTN-Network 'ENRAGE': *Random Geometry and Random Matrices: From Quantum Gravity to Econophysics* under Grant No. MRTN-CT-2004-005616 and the Deutsch-Französische Hochschule (DFH). The main part of the simulations was performed on the supercomputer JUMP of the John von Neumann Institute for Computing (NIC), Forschungszentrum Jülich under Grant No. hlz10.

References

1. A.P. Young (ed.), *Spin Glasses and Random Fields* (World Scientific, Singapore, 1997).
2. K.H. Fischer and J.A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, England, 1991).
3. M. Mezard, G. Parisi, and M.A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
4. K. Binder and A.P. Young, Rev. Mod. Phys. **58** (1986) 801.
5. M. Talagrand, C. R. Acad. Sci. Paris, Ser. I **337** (2003) 111.

6. G. Parisi, Phys. Rev. Lett. **43** (1979) 1754.
7. D. Sherrington and S. Kirkpartick, Phys. Rev. Lett. **35** (1975) 1792.
8. S. F. Edwards and P. W. Anderson, J. Phys. F: Metal Phys. **5** (1975) 965.
9. B.A. Berg and W. Janke, Phys. Rev. Lett. **80** (1998) 4771.
10. K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65** (1996) 1604.
11. B.A. Berg, A. Billoire, and W. Janke, Phys. Rev. B **61** (2000) 12143.
12. M. Palassini and S. Caracciolo, Phys. Rev. Lett. **82** (1999) 5128; H.G. Ballesteros, A. Cruz, L.A. Fernandez, V. Martin-Mayor, J. Pech, J.J. Ruiz-Lorenzo, A. Tarancon, P. Tellez, C.L. Ullod, and C. Ungil, Phys. Rev. B **62** (2000) 14237; M. Pleimling and I.A. Campbell, Phys. Rev. B **72** (2005) 184429.
13. B.A. Berg, Fields Inst. Commun. **26** (2000) 1.
14. W. Janke, Physica A **254** (1998) 164.
15. B.A. Berg, J. Stat. Phys. **82** (1996) 323.
16. W. Janke, *Histograms and All That*, in: *Computer Simulations of Surfaces and Interfaces*, NATO Science Series, II. Mathematics, Physics and Chemistry, Vol. **114**, eds. B. Dünweg, D.P. Landau, and A.I. Milchev (Kluwer, Dordrecht, 2003); p. 137.
17. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21** (1953) 1087.
18. E. Bittner and W. Janke, Europhys. Lett. **74** (2006) 195.
19. N.D. Mackenzie and A.P. Young, Phys. Rev. Lett. **49** (1982) 301; J. Phys. C: Solid State Phys. **16** (1983) 5321.
20. K. Nemoto, J. Phys. A: Math. Gen. **21** (1988) L287.
21. G.J. Rodgers and M.A. Moore, J. Phys. A: Math. Gen. **22** (1989) 1085.
22. D. Vertechi and M.A. Virasoro, J. Phys. (France) **50** (1989) 2325; Europhys. Lett. **12** (1990) 589.
23. S.G.W. Colborne, J. Phys. A: Math. Gen. **23** (1990) 4013.
24. H. Kinzelbach and H. Horner, Z. Phys. B **84** (1991) 95.
25. H. Takayama, H. Yoshino, and K. Hukushima, J. Phys. A: Math. Gen. **30** (1997) 3891.
26. A. Billoire and E. Marinari, J. Phys. A: Math. Gen. **34** (2001) L727.
27. L.A. Pastur and M.V. Shcherbina, J. Stat. Phys. **62** (1992) 1.
28. P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S.N. Coppersmith, Phys. Rev. Lett. **92** (2004) 097201.
29. E.J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
30. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed. (Krieger Publishing, Malibar, Florida, 1987).

Realistic Description of TTF-TCNQ – a Strongly Correlated Organic Metal

Andreas Dolfen and Erik Koch

Institut für Festkörperforschung, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {a.dolfen, e.koch}@fz-juelich.de

Understanding the physics of strongly correlated materials is one of the grand challenges in condensed-matter physics. Simple approximations such as the local density approximation fail, due to the importance of the Coulomb repulsion between localized electrons. Instead we have to resort to non-perturbative many-body techniques. Such calculations are, however, only feasible for quite small model systems. This means that the full Hamiltonian of a real material has to be approximated by a model Hamiltonian comprising only the most important electronic degrees of freedom, while the effect of all other electrons can merely be included in an average way in form of parameters. In this work we describe how to calculate those parameters for the one-dimensional organic metal TTF-TCNQ. Having constructed the Hamiltonian we calculate the ground state and dynamical properties with the Lanczos method. This method is limited by the available main memory. We show how to make efficient use of the memory and computational power of the massively parallel BlueGene/L system for such calculations. To gain high-resolution angular-resolved spectral functions we employ cluster perturbation theory (CPT) which helps identifying signatures of spin-charge separation also found experimentally in TTF-TCNQ. Increasing the nearest neighbour interaction is studied using a periodic version of CPT (kCPT).

1 Motivation

Essentially all of condensed matter physics is described by the non-relativistic Schrödinger equation $i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = H\left|\Psi\right\rangle$, with the Hamiltonian

$$H = -\sum_{\alpha=1}^{N_n}\frac{\vec{P}_\alpha^2}{2M_\alpha} - \sum_{j=1}^{N_e}\frac{\vec{p}_j^2}{2m} - \sum_{j=1}^{N_e}\sum_{\alpha=1}^{N_n}\frac{Z_\alpha e^2}{|\vec{r}_j-\vec{R}_\alpha|} + \sum_{j<k}^{N_e}\frac{e^2}{|\vec{r}_j-\vec{r}_k|} + \sum_{\alpha<\beta}^{N_n}\frac{Z_\alpha Z_\beta e^2}{|\vec{R}_\alpha-\vec{R}_\beta|}$$

where Z_α is the atomic number, M_α the mass, $\vec{R}_\alpha$ the position and $\vec{P}_\alpha$ the momentum of nucleus α. $\vec{p}_j$ and $\vec{r}_j$ denote the j^{th} electron's momentum and position and N_e, N_n the number of electrons and nuclei, respectively. To accurately describe materials of technological interest and design new ones with superior properties, all we have to do is solve this equation. There is, however, a severe problem which makes a brute-force approach to the many-body Schrödinger equation infeasible. To illustrate this, let us consider a single iron atom. With its 26 electrons the total electronic wave function depends on 26 times 3 spatial coordinates. Thus, even without spin, specifying the electronic wave function on a hypercubic grid with merely 10 points per coordinate, we would have to store 10^{78} numbers. This is impossible in practice: Even if we could store a number in a single hydrogen atom, the required memory would weight 10^{51} kg – far more than our home-galaxy, the milky way.

Still, the quantitative description of solids is not an entirely hopeless enterprise. Even though an exact treatment is a practical impossibility, there are successful approximations that work for wide classes of materials. The most prominent examples are approximations to density functional theory.[1] They effectively map the hard many-body problem to

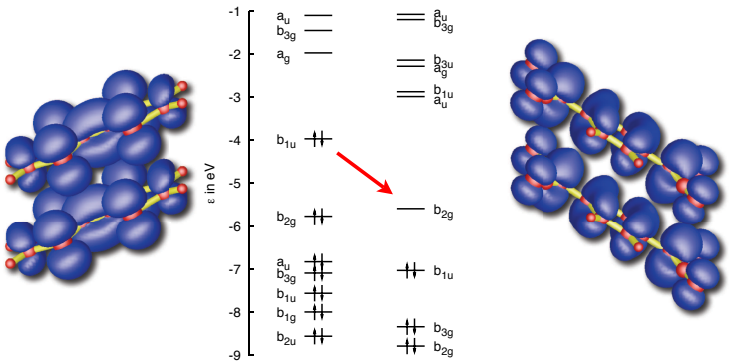


Figure 1. The molecular metal TTF-TCNQ. Centre: molecular levels of the isolated molecules; left: two TTF molecules with the electron density of their highest occupied molecular orbital (HOMO); right: TCNQ with the electron density of the lowest unoccupied molecular orbital (LUMO). The red arrow denotes the charge transfer of 0.6 electrons from the TTF-HOMO to the TCNQ-LUMO.

an effective single-particle problem that can be efficiently solved numerically. Essential to these approximations is that the Coulomb repulsion is described on a mean-field level. Such an approximation fails, however, to capture the physics in systems with strong correlations. In these systems the Coulomb repulsion between the electrons is so strong that the motion of a single electron depends on the position of all the others. The electrons thus lose their individuality and the single-electron picture breaks down. To accurately model this, we have to solve the many-electron problem exactly. Clearly we cannot do this for the full Hamiltonian. Instead, we consider a simplified Hamiltonian, which describes only those electrons that are essential to the correlation effects.[2] In this work we find the model Hamiltonian for the one-dimensional organic metal TTF-TCNQ and solve it numerically.

As shown in figure 1, TTF and TCNQ are stable molecules with completely filled molecular orbitals. The highest molecular orbital (HOMO) of TTF is, however, significantly higher in energy than the lowest unoccupied molecular orbital (LUMO) of TCNQ. Thus in a crystal of TTF and TCNQ, charge is transferred from the TTF-HOMO to the TCNQ-LUMO. This leads to partially filled bands and thus metallic behaviour. In the TTF-TCNQ crystal, like molecules are stacked on top of each other. Electrons can move along these stacks, while hopping between different stacks is extremely weak. Thus the material is quasi one-dimensional.

As pointed out above, we cannot treat all the electrons in the molecular solid. Instead, we focus our efforts on the most important electronic states: the partially filled TTF-HOMO and TCNQ-LUMO. The effects of the other electrons are included by considering their screening effects.[3] The simplest model Hamiltonian which captures both effects, the itinerancy of the electrons as well as the strong Coulomb interaction is the Hubbard model. We present a slightly extended Hubbard model, it reads

$$H = -\sum_{\sigma, i \neq j} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} + V \sum_{<ij>} n_i n_j \ . \qquad (1)$$

The first term gives the kinetic energy, where t_{ij} is the amplitude for an electron to hop from the molecule at site i to lattice site j. Note that hopping does not change the spin σ. The second and third terms represent the Coulomb repulsion between electrons in the same molecular orbital (second term) and in neighbouring molecular orbitals (third term). The values of U and V determine the strength of the Coulomb repulsion for two electrons in the same molecular orbital or neighbouring ones, respectively. In the next section we will address the problem of calculating those parameters.

Since the Hamiltonian does neither change the number of electrons nor their spin, we need to consider (1) only on Hilbert spaces with a fixed number of electrons of spin up $N_\uparrow$ and spin down $N_\downarrow$. For a finite system of L orbitals there are $\binom{L}{N_\sigma}$ different ways to arrange N_σ electrons of spin σ. Thus the dimension of the Hilbert space is given by $\binom{L}{N_\uparrow} \cdot \binom{L}{N_\downarrow}$. Even though we significantly simplified the problem, we still have to face the many-body problem: increasing system size, the dimension of the Hilbert space increases steeply. A system with 20 orbitals and 10 electrons of either spin already contains more than 34 billion ($34\,134\,779\,536$) different configurations. Storing a single many-body state for this system takes about 254 GB.

2 Calculation of Parameters for a Realistic Description

For a realistic description of TTF-TCNQ we employ all-electron DFT using the Perdew-Burke-Ernzerhof functional.[4] We start with the hopping matrix elements t along stacks of like molecules. Due to the relatively small overlap of the molecular wave functions the description in terms of the tight-binding model is a good approximation, for the same reason nearest neighbour hopping suffices. Similar to atomic energy levels splitting in bonding and anti-bonding levels when forming diatomic homonuclear molecules the molecular energy levels of isolated molecules split when two molecules approach one another. From this bonding and anti-bonding splitting for a pair of like molecules we obtain the absolute value of the hopping parameter. Let ν denote the molecular level for which we want to calculate the hopping parameter, here either the HOMO of TTF or the LUMO of TCNQ, and ε_ν its molecular energy. In a specific dimer with states $|\phi_\nu^x\rangle$, where $x = A, B$ distinguishes the two molecules, the tight-binding Hamiltonian reads

$$H_\nu^{\mathrm{TBA}} = \begin{pmatrix} \varepsilon_\nu & -t^\nu \\ -t^\nu & \varepsilon_\nu \end{pmatrix} . \tag{2}$$

Diagonalizing yields the symmetric/anti-symmetric state, i.e. $|\varphi_\nu^{s/a}\rangle = \frac{1}{\sqrt{2}} \left(\left|\phi_\nu^A\right\rangle \pm \left|\phi_\nu^B\right\rangle \right)$, with the corresponding energies, i.e. $|\epsilon_\nu^{s/a}\rangle = \varepsilon_\nu \mp t^\nu$. From this we obtain the absolute value of the hopping parameter by dividing the splitting $|\Delta\varepsilon^\nu| = 2|t^\nu|$ of the molecular energy levels by two. The sign of t^ν, however, is not directly accessible by this method. It can be derived from the symmetry of the dimer wavefunctions. If the bonding orbital (the lower one in energy) is symmetric ($\left|\varphi_\nu^{s/a}\right\rangle$) then t_ν is positive. And correspondingly if the lower one is antisymmetric, t is negative. For TTF-TCNQ we obtain: $t = -0.15$ eV for TTF and 0.18 eV for TCNQ.

The Coulomb parameters are harder to determine due to the screening processes inside the crystal. We start with the bare Coulomb integrals of the molecular orbital ν for two

	TTF	TCNQ		TTF	TCNQ
$U_{\rm bare}$	5.9	5.4	$V_{\rm bare}$	3.1	2.9
U_0	4.7	4.3	V_0	2.9	2.8
U	2	1.7	V	1	0.9

Table 1. Hubbard parameters for TTF-TCNQ. $U_{\rm bare}$ is the direct Coulomb integral, U_0 includes intra-molecular screening, and U is the screened on-site Coulomb term in the crystal whereas V denotes the screened nearest neighbour parameter. All energies are in eV.

molecules a relative distance of $\vec{l}$ apart, i.e.

$$V_{\rm bare}^{\nu,\vec{l}} = \int d^3\vec{r}d^3\vec{r}' \ \frac{\rho_\nu^0(\vec{r})\rho_\nu^{\vec{l}}(\vec{r}')}{|\vec{r}-\vec{r}'|} \ , \tag{3}$$

where $\rho_\nu^{\vec{l}}(\vec{r}) = |\phi_\nu^{\vec{l}}(\vec{r})|^2$ and $\phi_\nu^{\vec{l}}(\vec{r})$ is the wave function of orbital ν at position $\vec{l}$. Obviously the local Coulomb integral, the Hubbard-$U_{\rm bare}$, is given by $V_{\rm bare}^{\nu,\vec{0}} = U_{\rm bare}$. The bare parameters are, in general, too large since all screening effects are neglected. To calculate the on-site Coulomb parameter U_0 including intra-molecular screening we use (all-electron) density functional theory total energy calculations for different additional charges q on a single molecule. To obtain U_0 for HOMOs electrons are taken away and similarly for LUMOs electrons are added to the molecule. The Kohn-Sham DFT total energy consists of several contributions. It reads

$$E^{\rm total}[n] = \sum_{i=1} \varepsilon_i n_i - \int d^3\vec{r}\, V_{\rm xc} n(\vec{r}) - \frac{e^2}{2}\int d^3\vec{r}d^3\vec{r}' \ \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|} + E_{\rm xc}[n] + V_{\rm ions} \ , \tag{4}$$

where the ε_i are the Kohn-Sham eigenenergies and n_i the corresponding occupation numbers, $V_{\rm ions}$ denotes the ion-ion interaction.

The terms behave differently when subjected to a change in electron density. Charging up the molecule with q leads to a linear change in the occupation numbers n_ν for the orbital ν with energy ϵ_ν. The Hartree potential also has a linear contribution. This can be seen when substituting $\rho \to \rho + \delta\rho$, where $\delta\rho$ denotes the change due to the additional charge. Its contribution however is mainly quadratic. The effects of the exchange-correlation potential are usually small and therefore neglected. Thus, fitting the total energies for different additional charges to the function

$$E_{U_0^\nu}(q) = a_0 + a_1 q + U_0^\nu \frac{q^2}{2} \ , \tag{5}$$

yields the desired U_0 and a_0, a_1 as further fit parameters. a_0 should be about the energy for $E_{U_0^\nu}(0)$. a_1 captures the linear effects in q. The factor $1/2$ of the quadratic term stems from the factor $1/2$ in the Hartree potential in equation (4).

Similarly $V_0^{\nu,\vec{l}}$ is calculated. We consider two molecules with an additional charge of $q/2$ put on both of them. The total energy for different q is fitted to

$$E_{V_0^{\nu,\vec{l}}}(q) = 2E_{U_0^\nu}(q/2) + b_0 + b_1 q + V_0^{\nu,\vec{l}}\left(\frac{q}{2}\right)^2 \ , \tag{6}$$

yielding $V_0^{\nu,\vec{l}}$.

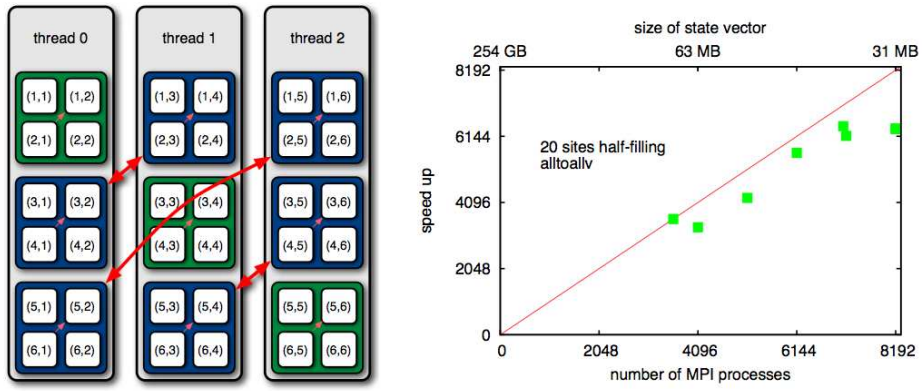


Figure 2. (left) Transpose operation that makes memory access thread-local when calculating the operation of the Hamiltonian on the state-vector. The communication (red arrows) is realized by a call to `MPI_alltoall`, which is very efficiently implemented on BlueGene/L. The small pink arrows indicate the local operations needed to complete the matrix-transpose;(right) Speed-up of our Lanczos code on IBM Blue Gene/L JUBL in CO mode for 20 sites half-filled Hubbard model.

To also include the inter-molecular screening contribution requires calculations of the energy of an infinite lattice of molecules. We employ an electrostatic approach and represent the molecules by their polarizabilities. The polarizability tensor of the isolated molecules is calculated with DFT, by evaluating the dipole moments in homogeneous external fields along the principal axes and extracting the linear response. Regarding the molecules as point-polarizabilities we can calculate the fully-screened parameters[3] with the distributed dipole-approach. All parameters are compiled in table 1.

3 Computational Aspects

The key ingredient of the Lanczos algorithm, our eigenvalue solver, is the sparse matrix vector multiplication. Already for quite small systems this operation takes most of the execution time, and with increasing the size of the many-body vector it dominates even more. Thus it will be in the focus of our parallelization efforts. On shared memory systems this matrix-vector multiplication is embarrassingly simple but we are restricted to relatively small memories. To use the memory that is needed to reduce finite size effects, we had to find an efficient distributed memory implementation.

The kinetic energy term of the Hamiltonian (1) has non-diagonal terms and therefore leads to non-local memory access patterns. To obtain an efficient distributed memory implementation we use a simple yet important observation: As pointed out above, the kinetic energy term conserves spin. Thus, performing the up-electron hopping takes only different up-hopping configurations into account while the down-electron configuration remains unchanged. If we group all up configurations for a fixed down configuration together in a single thread this hopping can thus be carried out locally: for a fixed index $i_\downarrow$, all $i_\uparrow$ configurations follow and can be stored in a thread. We see, that this basis can be naturally indexed by a tuple $(i_\downarrow, i_\uparrow)$ instead of a global index We can therefore equivalently regard

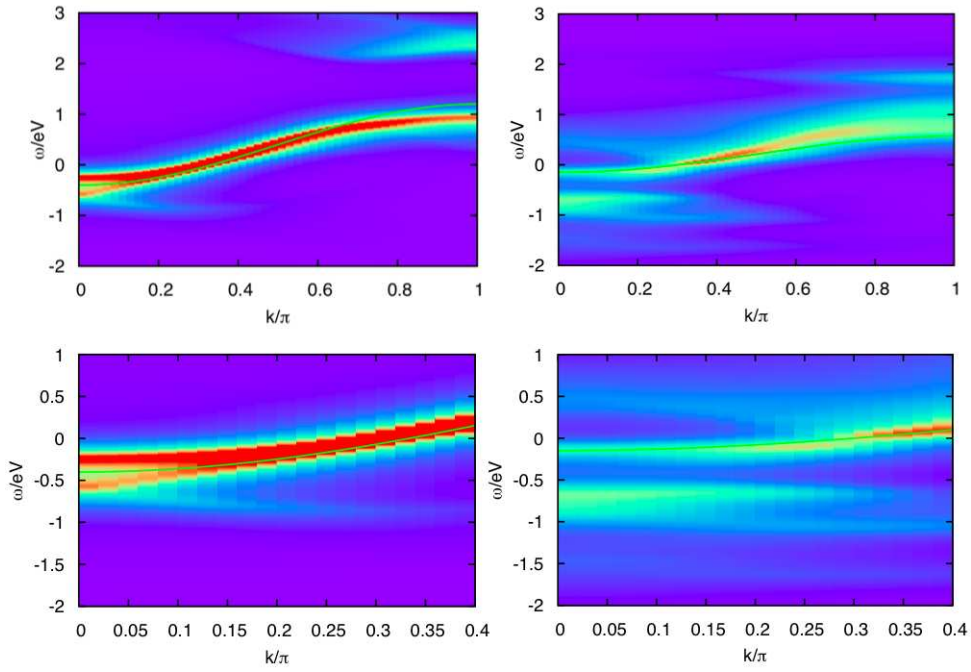


Figure 3. (first row) Angular-resolved spectral function obtained by CPT for a 20 sites TCNQ-like, i.e. six electrons of either spin, t-U Hubbard model with $U = 1.96$ eV, $t = 0.4$ eV (left) and t-U-V Hubbard model with $U = 1.7$ eV, $t = 0.18$ eV and $V = 0.9$ eV (right). The second row shows a magnification of the vicinity of the Γ point. In the t-U model we clearly observe signatures of spin-charge separation, whereas for the t-U-V model the lower branch is split. The green cosine shows the independent-particle band.

the vectors as matrices $v(i_\downarrow, i_\uparrow)$ with indices $i_\downarrow$ and $i_\uparrow$. Now it is easy to see that a matrix transpose reshuffles the data elements such that the down configurations are sequentially in memory and local to the thread. Therefore, the efficiency of the sparse matrix-vector multiplication rests on the performance of the matrix transpose operation. We implement it with `MPI_Alltoall`. This routine expects, however, the data packages which will be sent to a given process to be stored contiguously in memory. This does not apply to our case, since we would like to store the spin-down electron configurations sequentially in memory. Thus, the matrix is stored column wise. For `MPI_Alltoall` to work properly, we would have to bring the data elements in row-major order. This could be done by performing a local matrix transpose. The involved matrices are, however, in general rectangular, leading to expensive local-copy and reordering operations. We can avoid this by calling `MPI_Alltoall` for each column separately. After calling `MPI_Alltoall` for each column (red arrows in figure 2) only a local strided transposition has to be performed (small pink arrows) to obtain the fully transposed matrix or Lanczos vector.[5,6] The speed-up (figure 2) shows that collective communication is indeed very efficient on BlueGene/L.

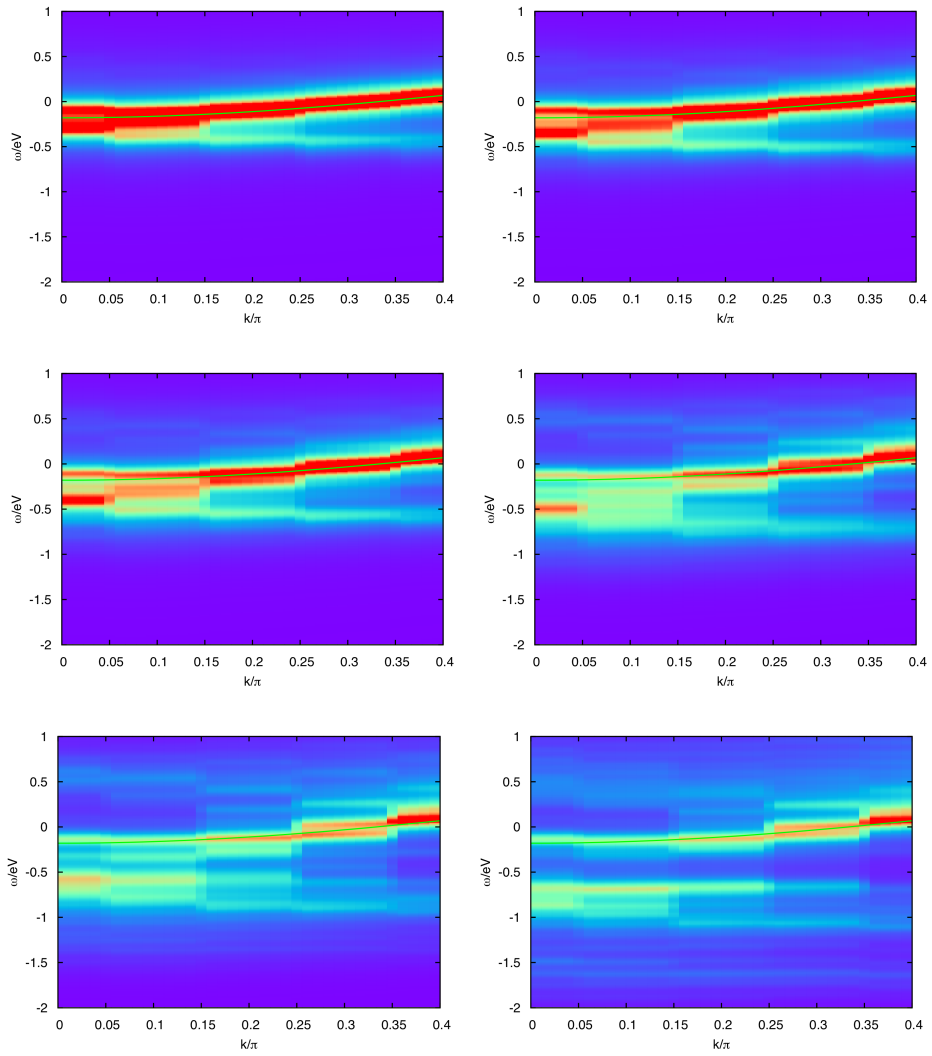


Figure 4. kCPT angular-resolved spectral functions close to the Γ point for 20 sites TCNQ-like Hubbard model with 6 electrons of either spin and different values of the nearest neighbour Coulomb repulsion $V = \{0.0, 0.1, 0.2, 0.4, 0.6, 0.9\}$eV ($U = 1.7$eV, $t = 0.18$eV). We clearly observe the increase in the splitting of the former holon and spinon branch with increasing values of V.

4 (k)Cluster Perturbation Theory and Spin-Charge Separation

Our parallel implementation of the Lanczos method enables us to efficiently calculate angular-resolved spectral functions for quite large systems. However, we still can have at most as many different momenta as we have sites. To resolve exciting physics like spin-charge separation we need, however, a higher resolution. A way to achieve this is cluster

perturbation theory (CPT).[7] The general idea is to solve a finite cluster with *open boundary conditions* exactly and then treat hopping between clusters in strong coupling perturbation theory, leading to an effectively infinite chain.

Figure 3 shows the angular-resolved spectral function for TCNQ in a CPT calculation for a 20 sites t-U Hubbard model. At the Γ-point we observe signatures of spin-charge separation: The electron dispersion splits into a holon and a spinon branch. These features are also observed in experiments. Usually the parameter set $U = 1.96$ eV, $t = 0.4$ eV has been used in the calculations since they fit the experiments fairly well. The parameter calculations, however, show that t should be smaller by a factor of more than two, but on the other hand the nearest neighbour interaction V should not be neglected since it is about half the value of U. We thus repeated the calculation with the realistic parameters derived above and observe that V effectively doubles the bandwidth. At the same time the spectral-weight of the holon-band spreads and the simple Luttinger liquid behaviour is lost.

To study this transition we need several calculations for different values of V. CPT calculations are, however, quite expensive. To generate the CPT plots about three BlueGene/L rack-days are needed: The calculation of the ground state is negligible and takes considerably less than half an hour on 2048 processors in VN mode on a BlueGene/L system. To calculate the Green's function for photoemission and inverse photoemission about 400 Green's functions each have to be calculated, where the former calculation takes a total of about 15 hours whereas the latter one takes about two days. We thus want to resort to a computationally less demanding, however, more finite-size effect prone method – the kCPT method. The key idea is to always keep translational symmetry by using periodic boundary conditions. Figure 4 shows how increasing V shifts apart the spectral features close to the Γ point.

References

1. W. Kohn, *Nobel Lecture: Electronic structure of matter: wave functions and density functionals*, Rev. Mod. Phys. **71**, 1253, 1999.
2. E. Koch and E. Pavarini, *Multiple Scales in Solid State Physics*, Proceedings of the Summer School on Multiscale Modeling and Simulations in Science 2007, Springer.
3. L. Cano-Cortés, A. Dolfen, J. Merino, J. Behler, B. Delley, K. Reuter, and E. Koch, *Coulomb parameters and photoemission for the molecular metal TTF-TCNQ*, Eur. Phys. J. B **56**, 173, 2007.
4. J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865, 1996.
5. A. Dolfen, *Massively parallel exact diagonalization of strongly correlated systems*, Diploma Thesis, RWTH Aachen, October 2006.
6. A. Dolfen, E. Pavarini, and E. Koch, *New Horizons for the Realistic Description of Materials with Strong Correlations*, Innovatives Supercomputing in Deutschland **4**, 16, Spring 2006.
7. D. Sénéchal, D. Perez, and M. Pioro-Ladrière, *Spectral Weight of the Hubbard Model through Cluster Perturbation Theory*, Phys. Rev. Lett. **84**, 522, 2000.
8. T. Maier, M. Jarrell, Th. Pruschke, M. Hettler, *Quantum Cluster Theories*, Rev. Mod. Phys **77**, 1027, 2005.

Soft Matter- and Nano-Systems: Computer Simulations

Florian Bürzle, Kerstin Franzrahe, Peter Henseler, Christine Schieback, Markus Dreher, Jörg Neder, Wolfram Quester, Daniel Mutter, Margit Schach, and Peter Nielaba

Physics Department, University of Konstanz, 78457 Konstanz, Germany
E-mail: {firstname.surname}@uni-konstanz.de

Soft matter systems have been investigated by Monte Carlo and Brownian Dynamics simulations. In particular the behaviour of two dimensional binary hard disk mixtures in external periodic potentials has been studied as well as the transport of colloids in micro-channels and the features of proteins in lipid bilayers. Ni nanocontacts have been analyzed by Molecular Dynamics simulations with respect to their conductance and structural properties under stretching. The properties of Si clusters in external fields have been computed by density functional methods, and static and dynamic properties of magnetic model systems by the Landau-Lifshitz-Gilbert equation. In the following sections an overview will be given on our recent results.

1 Two-Dimensional Model Colloids in External Periodic Fields

In monolayers on crystalline surfaces one can observe an intricate competition between effects due to the interaction of components within the layer and those with the underlying substrate.

Such complicated experimental systems can be modelled by two-dimensional colloidal systems. The interactions within the monolayer can be altered by changing the interaction potential of the colloids, while the shape and strength of the substrate potential can be modelled by external light fields. The advantage of the model system is, that via laser scanning microscopy direct access to the particle configurations is given. In this way it is possible to gain insight in the relative importance of the various possible physical processes that occur. From the theoretical point of view, even the relatively simple combination of a monodisperse system in a one-dimensional, spatially periodic light field shows a highly non-trivial phase behaviour as the amplitude of the external field is raised: Laser Induced Freezing (LIF) and Laser Induced Melting (LIM).

The interesting LIF- and LIM- effects have been studied in the HPC project by Monte Carlo simulations in two dimensions using commensurate[1,2] and incommensurate potentials. In addition, interesting ordering phenomena of two-dimensional colloidal crystals confined in strips of finite widths have been analysed by MC simulations[3,4].

In particular we explored[5] a hard disk system with commensurability ratio $p = \sqrt{3}a_s/(2\lambda) = 2$, where a_s is the mean distance between the disks and λ the period of the external potential. Three phases, the modulate liquid, the locked smectic and the locked floating solid have been observed, in agreement with other experimental[6] and analytical[7] studies. Various statistical quantities like order parameters, their cumulants and response functions, have been used to obtain a phase diagram for the transitions between these three phases.

For our analysis, we consider a system of hard disks with diameter σ. This system is

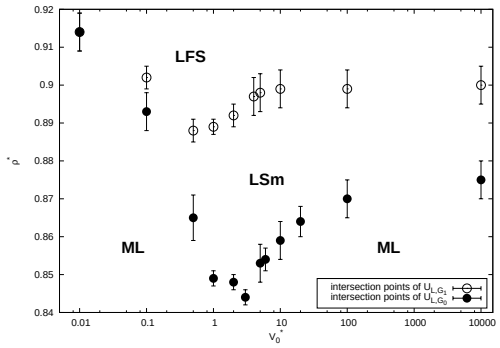


Figure 1. Phase diagram in the ρ^*/V_0^* plane. Transitions points have been obtained by considering order parameter cumulant intersection points.

subjected to an external potential

$$V(x,y) = V_0 \sin\left(\frac{2\pi}{\lambda}x\right) \tag{1}$$

Our system is characterized by the reduced density $\varrho^* = \varrho\sigma^2$ and the reduced potential strength $V_0^* = V_0/(k_B T)$, where k_B is the Boltzmann constant and T the temperature. For simplification, σ was set to unity in our simulation.

The resulting phase diagram shown in Fig. 1 was obtained by using the data from the cumulant intersection points. As the most important result, we see that the melting curves for both order parameters show a distinct remelting behaviour at higher V_0^*, as was expected by the theory of Radzihovsky *et al.*[7]. The melting curve from LFS to LSm (open circles) resembles those obtained by Strepp *et al.*[1] for $p = 1$ quite well. The other transition curve from ML to LSm (closed circles) shows that here the global minimum of the curve is slightly shifted to higher potential strengths. Also the minimum is located at considerable lower densities. Finally, it must be emphasized that at $V_0^* \to 0$, the different melting curves collapse into one single curve, as is expected for physical reasons.

In the studies of bi-disperse colloidal crystals our emphasis lies on the analysis of their structural and elastic properties. Monte Carlo simulations are an effective means for such analysis. We are interested in the dependence of these properties on the mixing ratio and size ratio of the components[8]. Monte Carlo simulations for hard disk mixtures with different diameter ratios σ_B/σ_A have been performed in the NPT- and NVT- ensemble in order to analyze the structural properties and phase transition parameters. Another point of interest is the phase behaviour of such systems in external, periodic light fields[9]. A bi-disperse hard disk mixture (mixing ratio 50% and diameter ratio $\sigma_B/\sigma_A = 0.414$) was exposed to an one dimensional, spatially periodic external light field. The wavelength of the external field was chosen to be commensurate to the square lattice, which yields the highest packing fraction for the given mixture. The commensurability ratio was set to $p = 2$. In these studies one has to distinguish between three cases: (I) only the smaller component interacts with the external field, (II) both components interact with the field and (III) only the larger component interacts with the field.

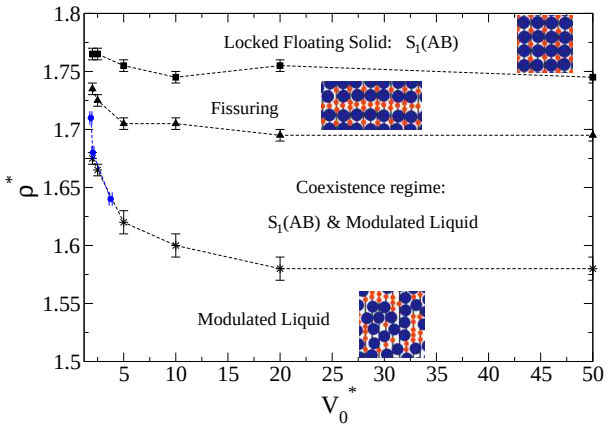


Figure 2. The $\varrho^* - V_0^*$ plane of the phase diagram of an equimolar binary mixture ($\sigma_B/\sigma_A = 0.414$) for the case, when only the smaller component interacts with the external light field.

Figure 2 shows[9] the phase diagram obtained for case (I). It was calculated by lowering the dimensionless number density $\varrho^* = \varrho\sigma_A^2$ and taking a commensurate path through phase space, meaning the wavelength of the external field $\lambda = 1/(\sqrt{2\varrho^*})$ changes with ϱ^*. Part of the phase diagram was obtained by raising the potential strength V_0^* at constant ϱ^*. Simulations carried out in an incommensurate setting, i.e. λ is kept constant independent of ϱ^*, intersect the phase diagram consistently.

At low external fields ($V_0 \leq 1.5$) we observe a laser induced coexistence of a triangular lattice of the larger component with a smaller component enriched binary fluid. In the field free case the system does not expose phase separation. This phase separation is driven by the attempts of the smaller components to form chains along the minima of the external field. Case (II) and (III) also show a laser induced de-mixing at low potential strengths. In these cases the larger component interacts directly with the external field. The phase separation is now also driven by the attempt of the larger component itself to align with the minima of the external field. The resulting coexisting monodisperse solid is a rhombic, commensurate lattice.

At higher external fields as shown[9] in figure 2 case(I), i.e. only the smaller component interacts with the external field, exhibits a laser induced freezing transition into the the commensurate $S_1(AB)$ square lattice. Depending on the overall density ϱ^* the $S_1(AB)$ locked floating solid is either in the one phase regime or coexists with an equimolar binary fluid. These two regimes are separated by a fissuring region, in which the smaller component is free to move perpendicular to the minima of the external field.

2 Transport of Colloids in Micro-Channels

We conducted Brownian dynamics (BD) simulations of a two-dimensional microchannel[10] setup in order to investigate the flow behaviour of the colloidal particles within the channel systematically for various parameter values of constant driving force, overall particle density, and channel width. The pair interaction $V(r) = (\mu_0/4\pi)M^2/r^3$ (M is the dipole

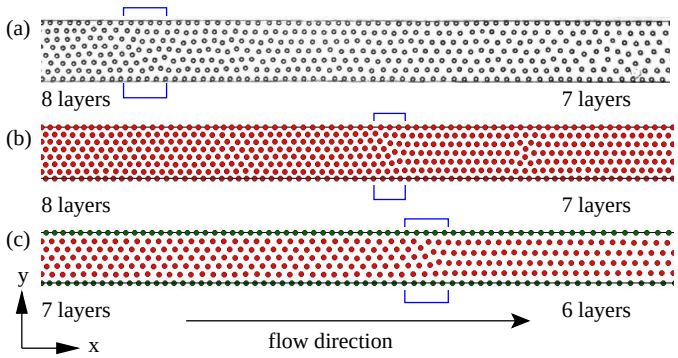


Figure 3. (a) Video microscopy snapshot of colloidal particles moving along the lithographically defined channel. (b) Simulation snapshots for a channel ($692 \times 60\mu m$, $\Gamma \approx 2.5$) with ideal hard walls ($573.3 \times 45\mu m$, $\Gamma = 115$), (c) the same as in (b) with the particles at the walls (marked green) kept fixed ($573.3 \times 45\mu m$, $\Gamma = 902$). The blue rectangles mark the layer transition region.

moment) is purely repulsive and can be characterized by the dimensionless interaction strength $\Gamma = \mu_0 M^2 \rho^{2/3}/(4\pi k_B T)$.

Particles are confined to the channel by ideal elastic hard walls in y-direction and at $x = 0$ (channel entrance). Also we performed simulations with the particles at the wall kept fixed. The channel end is realized as an open boundary. To keep the overall number density in the channel fixed, every time a particle leaves the end of the channel a new particle is inserted at a random position (avoiding particle overlaps) within the first 10% of the channel, acting as a reservoir. A cutoff of 10σ was used along with a Verlet next neighbour list[11]. Checks of particle overlaps are included in the simulation, but for all ordered systems we never found two overlapping particles.

A typical snapshot from the experiment[10] of the particles moving along the channel is shown in Fig. 3(a). Similar snapshots we get from simulations[10] with co-moving (Fig. 3(b)) and fixed boundary particles (Fig. 3(c)), i.e., the velocity is kept to zero for the particles at the channel wall. In most regions of the channel the particles are placed in a quasi-crystalline order. This behaviour is due to the strength of the particle interactions caused by the external magnetic field (high Γ-values), which leads to quasi-crystalline behaviour in unbounded systems as well. The formation of this order naturally gives rise to the formation of layers in the motion of the particles along the channel. A similar layering phenomenon has been observed in channels under equilibrium condition[12]. Additionally to this layer formation we observe, both in experiment and in simulation, a decrease of the number of layers in the direction of motion. In between both regions therefore a region exists in which the particles cannot be well-ordered. This region is called the layer-reduction zone. In Fig. 3 these regions have been marked.

The reduction of the number of layers originates from a density gradient along the channel. The local particle density inside the channel is shown in Fig. 4(a) and (b) together with the particle separations in x- and y-directions. In the experiment (Fig. 4(a)) the density decreases monotonically along the direction of the motion of the particles by about 20%. The average density in the channel shows fluctuations on the order of 10% as a function of time. The total increase in density, however, is less than 3% during the total time of

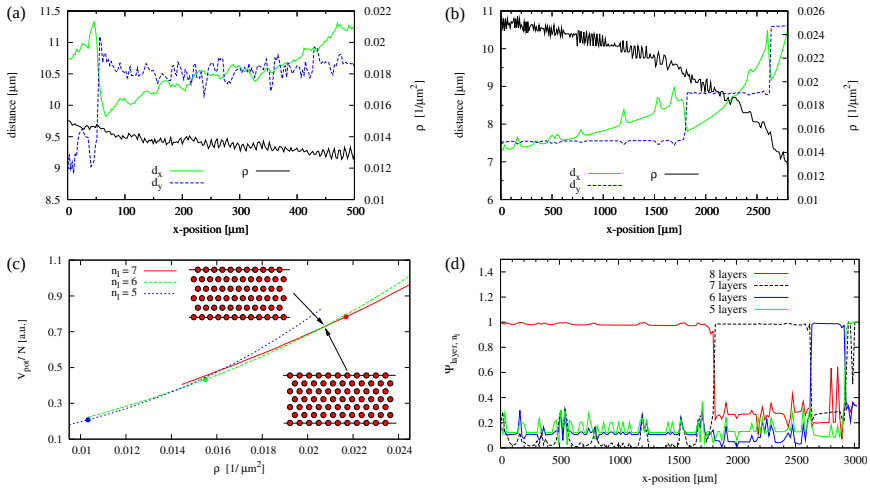


Figure 4. Local lattice constants d_x and d_y and local particle density (a) in the experiment and (b) in the BD simulation. The results are obtained for the systems of Fig. 3(a) and (b) respectively. (c) Potential energies per particle of different layer configurations as a function of the particle density. The dots mark the perfect triangular lattices for 5, 6 and 7 layers. Also shown are parts of the configurations with 7 and 6 layers at the intersection point. (d) Plots of the layer order parameter for the configuration snapshot of Fig. 3(b).

the experiment. We therefore argue that the density gradient is formed in a quasi-static situation. This argument is confirmed by results of BD simulations (Fig. 4(b)), where the corresponding decrease of the particle density is observed.

3 Conductance and Structural Properties of Ni Nanocontacts

During the last years a lot of attention has been devoted to the analysis of contacts of magnetic materials. In these nanowires the spin degeneracy is lifted, which can potentially lead to interesting spin-related phenomena in the transport properties.

Here we address the issue of the conductance quantization and the spin polarization of the current of Ni contacts. We have combined classical molecular dynamics simulations of the breaking of nanocontacts with conductance calculations based on a tight-binding model.[13,14] For Ni we have applied our method to a Hamiltonian with spin-dependent matrix elements.[15]

We analyzed the evolution of the conductance during the formation of a Ni dimer structure, which is the most common geometry in the last stages of the breaking process. In addition to the evolution of the conductance and transmission eigenchannels for both spin components separately, we computed the MCS radius, the strain force, and the spin polarization of the current, which is defined as $P = (G^{\uparrow} - G^{\downarrow})/(G^{\uparrow} + G^{\downarrow}) \times 100\%$, where G^{σ} is the conductance of the spin component σ. Here, spin up ($\sigma = \uparrow$) means majority spins and spin down ($\sigma = \downarrow$) minority spins. In the last stages of the stretching the conductance for the majority spins lies below $1.2e^2/h$ and is dominated by a single channel, while for the minority spin there are still up to four open channels and the conductance is close to $2e^2/h$, adding up to a total conductance of around 1.2–$1.6G_0$.

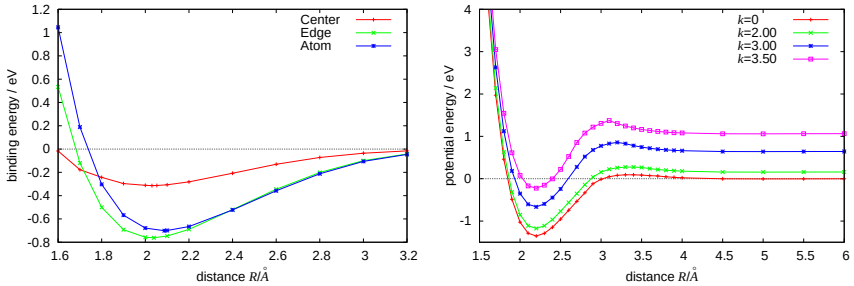


Figure 5. Left: Binding energy of Si atoms as function of distance over certain sites of a graphite surface Right: Potential energy for two Si_4 clusters as function of distance in an external parabolic potential.

For ferromagnetic Ni, we have shown that the contacts behave as a mixture of a noble metal (such as Ag) and a transition metal (such as Pt). While the $4s$ orbitals play the main role for the transport of the majority-spin electrons, the conduction of the minority-spin electrons is controlled by the partially occupied $3d$ orbitals. This follows from the position of the Fermi energy, which lies in the s band for the majority spins and in the d bands for the minority spins. Our results indicate the absence of any conductance quantization, and show how the spin polarization of the current evolves from negative values in thick contacts to even positive values in the tunneling regime after rupture of the contact.

4 Si_n Clusters, Magnetic Model Systems and Membranes under Tension

We computed the properties of selected Si_n clusters approaching each other and the effect of external fields and surfaces[17,18] by DFT methods[19].

One example shown here is the approach of Si-atoms to a graphite surface. The resulting binding energy as function of distance over certain sites[16] is shown in Fig. 5. It follows that the site above the C-C bonds is energetically preferred. Another example[18] is the effect of an external parabolic potential, $V(x, y, z) = k^2y^2$ on the potential energy of two Si_4 clusters as a function of distance, s. Fig. 5. Two nearest atoms of the two clusters are fixed at a distance R, the other atoms are allowed to move freely. For increasing k values the potential energy develops a fusion barrier at a distance of about 3 Å.

Using the Heisenberg Hamiltonian and the Landau-Lifshitz-Gilbert equation interesting insight into the behaviour of domain walls in confined geometry at finite temperature has been computed[20–22].

The dynamics of the spin reversal processes in systems with moments attached to caps of (colloidal) spheres have been computed[22] for different system sizes and external magnetic fields. Fig. 6 shows the hysteresis effect of external magnetic fields at angle ϑ relative to the x-axis.

Lipid bilayers and incorporated proteins form biological membranes. These barriers define the inside and the outside of a cell and are indispensable for live. Usually the microscopic surface tension of membranes is small or vanishes altogether. One aim of our work[23] is to study the effect of an applied tension to a model bilayer, unsing the lipid

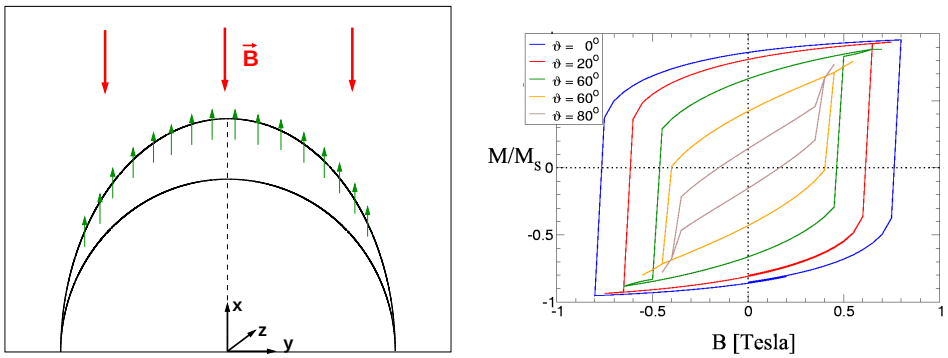


Figure 6. Left: Sketch of the geometry; Right: Hysteresis for different angles ϑ relative to the x-axis. The caps have a diameter of 48 nm and a height of 12 nm.

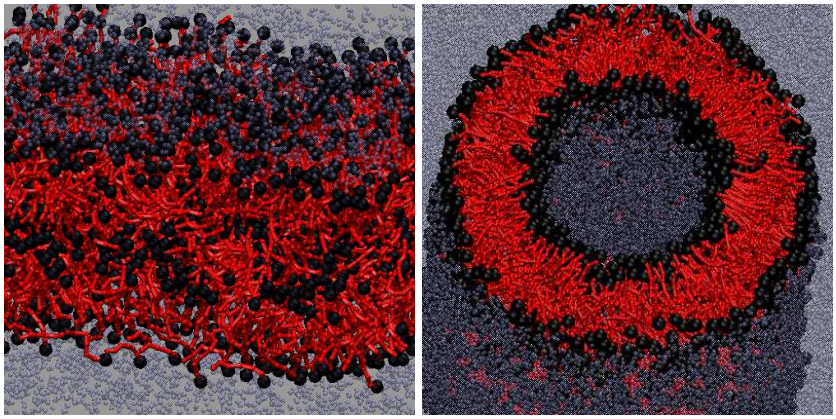


Figure 7. Left: Snapshot of a double bilayer configuration; Right: Tether consisting of 4,800 lipids

model of Ref.[24,26]. Does this tension cause a change in the behaviour of incorporated model proteins, e.g. lead to an increasing lipid mediated attraction or repulsion between two proteins?

Other points of interest in our project are the examination of membrane multi layers (Fig. 7) and tethers (Fig. 7) within and without an applied tension, respectively. These configurations[23] require high computing capacities due to the system size. The shown tether, e. g., consists of 4800 lipids and more than 90,000 solvent particles. The program has been parallelized and technical details about that can be found in Ref.[25].

Acknowledgments

We gratefully acknowledge useful discussions with K. Binder, C. Cuevas, F. Pauly, F. Schmid, S. Sengupta, B. West, support from the SFB 513, the SFB TR6 and the SFB 767 and granting of computer time from the NIC.

References

1. W. Strepp, S. Sengupta, P. Nielaba, Phys. Rev. **E63**, 046106 (2001).
2. P. Nielaba, W. Strepp, S. Sengupta, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, edited by M. Ferrario, G. Ciccotti, K. Binder, Springer, Berlin, pp.163 (2006).
3. A. Ricci, P. Nielaba, S. Sengupta, K. Binder, Phys. Rev. **E74**, 010404 (R) (2006).
4. A. Ricci, P. Nielaba, S. Sengupta, K. Binder, Phys. Rev. **E75**, 011405 (2007).
5. F. Bürzle, P. Nielaba, Phs. Rev. **E76**, 051112 (2007).
6. J. Baumgartl, M. Brunner, C. Bechinger, Phys. Rev. Lett. **93**, 168301 (2004).
7. E. Frey, D.R. Nelson, L. Radzihovsky, Phys. Rev. Lett. **83**, 2977 (1999); L. Radzihovsky, E. Frey, D.R. Nelson, Phys. Rev. **E63**, 031503 (2001).
8. K. Franzrahe, P. Henseler, A. Ricci, W. Strepp, S. Sengupta, M. Dreher, Chr. Kircher, M. Lohrer, W. Quester, K. Binder, P. Nielaba, Comp. Phys. Commun. **169**, 197 (2005).
9. K. Franzrahe, P. Nielaba, Phys. Rev. **E**, in press.
10. M. Köppl, P. Henseler, A. Erbe, P. Nielaba, and P. Leiderer. Phys. Rev. Lett. **97**, 208302 (2006).
11. M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids* (Oxford Science Publications, 1987).
12. R.Haghgooie, C.Li, P.Doyle, Langmuir**22**, 3601 (2006).
13. F. Pauly, M. Dreher, J.K. Viljas, M. Häfner, J.C. Cuevas, and P. Nielaba, Phys. Rev. B **74**, 235106 (2006).
14. M. Dreher, F. Pauly, J. Heurich, J.C. Cuevas, E. Scheer, and P. Nielaba, Phys. Rev. B **72**, 075435 (2005).
15. See parameters ni_ferro_par at `http://cst-www.nrl.navy.mil/bind/ni.html`; M.I. Haftel, N. Bernstein, M.J. Mehl, and D.A. Papaconstantopoulos, Phys. Rev. B **70**, 125419 (2004).
16. W. Quester, Dissertation, U. Konstanz (in preparation).
17. F. von Gynz-Rekowski, W. Quester, R. Dietsche, Dong Chan Lim, N. Bertram, T. Fischer, G. Ganteför, M. Schach, P. Nielaba, Young Dok Kim, Eur. Phys. J. **D**, in press.
18. M. Schach, Diplomarbeit, U. Konstanz (2007).
19. CPMD. Copyright IBM Corp 1990–2001, Copyright MPI für Festkörperforschung Stuttgart 1997–2004. `http://www.cpmd.org/`.
20. C. Schieback, M. Kläui, U. Nowak, U. Rüdiger, P. Nielaba, Eur. J. **B**, in press.
21. D. Backes, C. Schieback, M. Kläui, F. Junginger, H. Ehrke, P. Nielaba, U. Rüdiger, L.J. Heyderman, C.S. Chen, T. Kasama, R.E. Dunin-Borkowski, C.A.F. Vaz, J.A.C. Bland, Appl. Phys. Lett. **91**, 112502 (2007).
22. D. Mutter, Diplomarbeit, U. Konstanz (2007).
23. J. Neder, Dissertation, U. Konstanz (in preparation).
24. O. Lenz and F. Schmid, J. Mol. Liq. **117**, 147 (2005).
25. F. Schmid, D. Düchs, O. Lenz, B. West, Comp. Phys. Commun. **177**, 168 (2007).
26. O. Lenz and F. Schmid, Phys. Rev. Lett. **98**, 058104 (2007).

Computational Soft Matter Science

Kurt Kremer

Max Planck Institute for Polymer Research
55021 Mainz, Germany
E-mail: kremer@mpip-mainz.mpg.de

Synthetic polymers, bio membranes, proteins, colloids and many other systems are characterized by characteristically low modulii and therefore are also termed "soft matter". In turn "soft matter" means, that thermal fluctuations play an important role for the structure and properties of these materials and consequently conformational and structural fluctuations have to be considered. The material properties are intimately linked to both, mesoscopic generic aspects, such as molecular architecture (chain molecules, branched molecules etc.) and details of the chemical building unit. Since both aspects play an equally important role, models of both kinds are treated extensively within theoretical studies and nowadays in many cases these scales are also directly linked to each other. On the other hand, the complicated nature of the structure units combined with the above mentioned fluctuations makes it almost impossible up to some very specific exceptions to arrive at truly quantitative predictions on the basis of analytic theory only. Because of that computer simulations have already for many years played a very significant role in the theoretical study of soft matter. In this context, the NIC was instrumental for many groups throughout Germany to provide CPU power which fits the specific needs of this community. This is also reflected by the contributions to the NIC Symposium Volume 2008, which reflects this range of this theoretical description from mesoscopic simple generic models all the way to poly-atomistic descriptions. The present chapter of soft matter simulations contains three rather different contributions.

Free Energies of Self-Assembling Systems

The first contribution by Müller and Daoulas deals with the problem of self-assembly in soft matter. Different polymeric systems are usually very difficult to mix. Because of the high conformational entropy which to a first order is the same in the mixed and the phase-separated state only very small unfavourable interactions (of the order of the thermal energy over molecular size) are sufficient to drive phase segregation. This is a characteristic problem of high technical and simultaneously of theoretical interest. Combining different polymeric structures into one macromolecule then leads to very interesting morphologies with nano- or microscopic patterns due to the tendency of self-aggregation of chemically identical parts. However, to find these different phases and to estimate the free energy differences between them is quite difficult. In the present approach the SCMF (Self Consistent Mean Field) method is known to be a numerically very efficient method to study phase transitions and dense polymeric systems and is extended to provide free energy differences between different morphologies. The idea is to couple the density field of the macromolecular system to an ordering field which one can slowly switch on or off. By

this effectively a thermodynamic integration can be performed in order to obtain the free energies of different morphologies. The present example of lamellar diblock-copolymers displays the potential power of the approach. The method is easily parallelized which will allow the authors to study much more complicated and larger systems with the new machines recently arrived in Jülich.

Molecular Dynamics Simulations of PVA/Pt(111) Interfaces

While the first one was a study on the mesoscopic generic level, the second contribution by Tarmyshov and Müller-Plathe deals with very local atomistic properties of small macromolecules close to a metal surface. They study a melt of PVA (poly(vinyl alcohol)) on platinum Pt(111) and compare this to an earlier study of liquid isopropanol on the same platinum surface. While the isopropanol molecule close to the surface becomes weakly bonded to the platinum via the hydroxyl oxygen this is somewhat restricted in the case of PVA. Here, the steric properties of the chain molecule come into play and prevent a perfect link to the surface as it was found for isopropanol. Still, the PVA molecules coat the surface and form a rather compact layer, separating the surface linked molecules from the other bulk molecules.

Effective Protein Interactions in a Coarse-Grained Model for Lipid Membranes

Again, a completely different problem is addressed in the contribution by West and Schmid. Biological membranes are prototypical soft matter systems, and here especially the properties and function of membrane embedded proteins is of high interest to both, biologists as well as physicists. However, studies on that problem by atomistic simulations up to some very short time runs are beyond the capabilities even of the newest generation of supercomputers. Because of that scientists constructed reduced models which are as simple as possible in order to provide a realistic spectrum of the membrane fluctuations. West and Schmid study such a generic model where the lipids of the membrane are replaced by a simple sixmers and the embedded proteins are treated as simple cylindrical objects. The solvent is treated as an ideal gas of “head groups” of the simplified lipid molecules. By this the authors study the interaction of two embedded cylinders and, what is especially important, the coupling of this interaction to the capillary wave spectrum of the membranes. Besides extending these studies it is one of the challenges for the future to link such kind of computations to a more detailed understanding of membrane properties.

Outlook

Beyond the here described three projects the NIC provided CPU time grants for a number of further soft matter projects also including colloid physics and chemistry.

All the presented problems significantly benefitted from the CPU power of the Jülich supercomputers and are especially suitable for moderately parallel computing. This is the case because system sizes typically reach up to 1 Mio. particles and then have to be followed for a long time.

Accurate Measurement of Free Energies of Self-Assembling Systems by Computer Simulation

Marcus Müller and Kostas Ch. Daoulas

Institut für Theoretische Physik, Georg-August Universität
Friedrich-Hund-Platz 1, D-37077 Göttingen, Germany
E-mail: mmueller@theorie.physik.uni-goettingen.de

A method for calculating free energy differences between disordered and ordered phases of self-assembling systems is discussed. Applying an external, ordering field, we impose a predefined structure onto the fluid in the disordered phase. The structure in the presence of the external, ordering field closely mimics the structure of the ordered phase (in the absence of an ordering field). Subsequently, we gradually switch off the external, ordering field and, in turn, increase the control parameter that drives the self-assembly. The free energy difference along this reversible path connecting the disordered and the ordered state is obtained via thermodynamic integration or expanded ensemble simulation techniques in conjunction with successive umbrella sampling.

1 Introduction

Computer simulations provide accurate information about the statistical mechanics and thermodynamics without the need to invoke approximations that are often required to make progress in analytical calculations. The measurement of free energies, however, is a challenge because free energy differences cannot simply be expressed as functions of the particles' coordinates. Special simulation techniques have to be devised in order to extract free energy differences from particle simulations.[1]

In this report, we present a method that enables us to accurately calculate the free energy differences between self-assembled morphologies in amphiphilic systems. This is a computationally difficult problem for two reasons: (i) Due to the mismatch of the periodicity of the self-assembled morphology with the size of the simulation cell, there are strong finite-size effects.[2] Calculations with variable box shape can mitigate this problem,[3,4] and we do not consider this important aspect further here. (ii) Another problem stems from the absence of a well-ordered reference state – the analog of a crystal – for which the free energy can be accurately calculated. If the absolute free energy of such a reference state in the ordered phase was known, one could utilize thermodynamic integration to calculate the free energy at different state points (e.g. temperatures) and accurately map out the phase diagram without relying on the observation of hysteresis.

It is interesting to draw a comparison between self-assembly in soft matter (e.g., lamellar ordering in a diblock copolymer melt) and crystallization in simple, hard condensed matter systems (e.g., a Lennard-Jones crystal). The local difference of volume fraction, $\phi(\mathbf{r})$, of the two species of the amphiphilic system plays a similar role as the density of a hard crystal. Its dominant Fourier mode is the order parameter of the transition. In a well-ordered, hard crystal, each particle fluctuates little around its corresponding crystal lattice position. Thus, the system resembles an Einstein crystal, in which non-interacting particles are tethered by harmonic springs to their ideal lattice positions. Frenkel and Ladd

have used thermodynamic integration from this Einstein crystal to the well-ordered solid for calculating the absolute free energy of a hard crystal.[5]

In a self-assembling soft matter system, the composition also fluctuates little around the ideally ordered value, however, the molecules are in a liquid state, i.e. they diffuse and are not "tethered" to ideal positions. Therefore, there exists no well-defined reference state and previous simulation techniques for calculating the absolute free energy of hard crystals do not straightforwardly carry over to self-assembling soft matter systems.

In fact, even calculating the free energy of a homogeneous melt without repulsion between the two different monomeric species (i.e., $\chi N = 0$), which would be the analog of an ideal gas in a simple, hard condensed matter system, is a formidable task. In a liquid, the (bonded and non-bonded) interactions of a segment with its surroundings are of the order k_BT, where k_B and T are Boltzmann's constant and temperature, respectively. Thus, the free energy per molecule is proportional to k_BTN where N denotes the number of segments (or coarse-grained interaction centres) each molecule is comprised of. In order to accurately determine the location of phase boundaries, calculate the free energy costs of defects or grain boundaries, and assess the thermodynamic stability of morphologies, one needs to know the free energy per molecule with an accuracy of $\mathcal{O}(10^{-3}k_BT)$. Therefore the absolute free energy would be required to be known with a precision of the order $10^{-3}/N \approx 10^{-5}$. The free energy difference between the disordered phase and a self-assembled structure, however, is only of the order $\chi N k_BT \sim k_BT$. Thus, it is advantageous to directly calculate the free energy difference between the disordered and ordered state rather than to determine it as a difference of two large absolute free energies.

Here we present a general thermodynamic integration scheme that enables us to calculate the free energy *difference* between a disordered and a spatially structured phase. We illustrate the method by studying a symmetric diblock copolymer melt. We convert the disordered melt ($\chi_{\rm init}N < \chi_{\rm ODT}N$) into the spatially structured phase ($\chi_{\rm final}N > \chi_{\rm ODT}N$) via a *reversible path*. Our method is inspired by the work of Sheu, Mou and Lovett calculating the absolute free energy of a Lennard-Jones solid[6]. Related reversible integration paths between a solid and a liquid have been used by others[7,8] for calculating the free energy difference between a solid and a liquid.

First, as illustrated in Fig. 1, we structure the homogeneous, disordered melt at low incompatibility, $\chi_{\rm init}N$, between the two species by applying an external, ordering field, $h(\mathbf{r})$, conjugated to the order parameter of the transition. Ref. 6 demonstrated that the optimal choice of this ordering field, h, is such that the order parameter at $\chi_{\rm init}N$ in the presence of the ordering field closely mimics the order parameter at the final self-assembled state, $\chi_{\rm final}N$, in the absence of the ordering field. The structure formation in response to the ordering field is completely gradual and free of any thermodynamic singularities. Second, we gradually reduce the ordering field and, in turn, increase the incompatibility between the two species. Optimally, the spatial order parameter does not vary along this path. The absence of abrupt changes in the order parameter, $\phi(\mathbf{r})$, or any other quantities indicates the lack of thermodynamical singularities along the second branch of the path. Thus, one turns a disordered system into a spatially ordered one without passing through a first-order transition.[6]

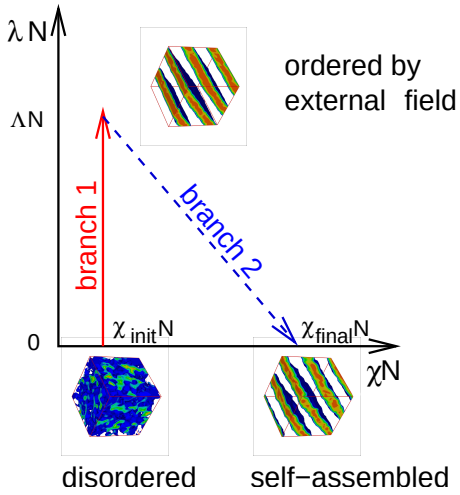


Figure 1. Sketch of the reversible path that connects the homogeneous, disordered state, the externally ordered and the self-assembled state. Configuration snapshots of a symmetric diblock melt illustrate the different states. In the snapshots three-dimensional contour plots of the composition are shown. The B-rich component is removed for clarity and the interface between the different components is coloured blue. $\chi_{\rm init}N = 0$, $\chi_{\rm final}N = 20$, and the maximal strength of the ordering field is $\Lambda N = 20$. The linear extension of the simulation cell is $L = 7.77R_{eo}$ and the lamellar spacing is $L_o = L/(2\sqrt{2}) = 1.686R_{eo}$.

2 Model and Technique

Our thermodynamic integration scheme can be applied to different models including coarse-grained, particle-based models of amphiphilic systems and membranes[9] as well as field-theoretic representations[10]. It can be implemented in Monte Carlo or Molecular Dynamics or DPD simulations, as well as Single-Chain-in-Mean-Field (SCMF) simulations,[11,12] field-theoretic simulations[10], and external potential dynamics[13,14] or dynamic density functional theory.[15]

In the following, we consider a liquid of n diblock copolymer molecules in a volume, V. Let the order parameter, $\phi(\mathbf{r})$, denote the difference of the local volume fractions of the two components, and χN the incompatibility of the two components per molecule. The energy of the system takes the form $\mathcal{H} = \mathcal{H}_{\rm liq} + \mathcal{H}_{\rm ord} + \mathcal{H}_{\rm ext}$.

$$\mathcal{H}_{\rm liq} = \mathcal{H}_{\rm b} + \mathcal{H}_{\rm nb} \quad \text{with} \quad \frac{\mathcal{H}_{\rm b}}{k_B T} = \sum_{i=1}^{n} \sum_{s=1}^{N} \frac{3(N-1)}{2{R_{eo}}^2} \left[\mathbf{r}_i(s) - \mathbf{r}_i(s+1)\right]^2 \qquad (1)$$

The bonded, intramolecular interactions, $\mathcal{H}_{\rm b}$, take the form of a discretized Edwards-Hamiltonian. The shape of a flexible macromolecule is Gaussian and characterized by its mean squared end-to-end distance, R_{eo}. We discretize the contour of the symmetric diblock copolymer into $N = 16 + 16$ coarse-grained segments. The coarse-grained parameter of relevance for matching our simulation data to experiments is the invariant degree of polymerization, $\bar{\mathcal{N}} = (nR_{eo}^3/N)^2 = 14\,884$. This value lies in the typical experimental

range, and we achieve this large value not by using a particularly fine discretization, N, along the molecules but rather by increasing the polymer density, n/V.

We calculate the coarse-grained compositions from the microscopic molecular conformations, $\{\mathbf{R}_i\}$, using a simple cubic grid with mesh size, ΔL, i.e., the grid contains $N_{\rm cells} = V/\Delta L^3$ cells. $\Delta L = 0.19875 R_{eo}$ provides sufficient spatial resolution. The local composition, $\phi_A(\mathbf{j})$, at a grid point, $\mathbf{r}_\mathbf{j} = j_x\hat{\mathbf{x}} + j_y\hat{\mathbf{y}} + j_z\hat{\mathbf{z}}$ (j_x, j_y, j_z, being multiples of ΔL, index the grid point and $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ denotes the unit cell vectors of the simple cubic lattice, respectively) is given by $\phi_A(\mathbf{j}) = \frac{V}{nN\Delta L^3}\sum_{i=1}^{n}\sum_{s=1}^{fN} m(\mathbf{r}_i(s) - \mathbf{r}_\mathbf{j})$, and a similar expression holds for the local volume fraction of the B component. The assignment function, m, characterizes the mapping of the continuous, off-lattice segment coordinates onto the coarse-grained grid and we utilize a linear particle-mesh extrapolation.[16]

The small compressibility of a dense polymer liquid is captured by the non-bonded interactions of Helfand-type: $\frac{\mathcal{H}_{\rm nb}}{nk_BT} = \frac{\kappa_o N}{2}\frac{1}{N_{\rm cells}}\sum_\mathbf{j}\left[\phi_A(\mathbf{j}) + \phi_B(\mathbf{j}) - 1\right]^2$. The value $\kappa_o N = 50$ is sufficient to suppress fluctuations of the total density on the length scale of a small fraction of R_{eo}.

The incompatibility between unlike segments, A and B, is described by the most symmetric choice: $\frac{\mathcal{H}_{\rm ord}}{nk_BT} = -\frac{\chi N}{N_{\rm cells}}\sum_\mathbf{j}\phi^2(\mathbf{j})$ with $\phi(\mathbf{j}) = \frac{\phi_A(\mathbf{j}) - \phi_B(\mathbf{j})}{2}$. In addition, we apply an external, ordering field, $h(\mathbf{r}) = -\lambda N f_{\rm ext}(\mathbf{r})$, that linearly couples to the order parameter. It results in an energy contribution of the form $\frac{\mathcal{H}_{\rm ext}}{nk_BT} = -\frac{\lambda N}{V}\int_V \mathrm{d}^3\mathbf{r}\ f_{\rm ext}(\mathbf{r})\phi(\mathbf{r})$ where λN characterizes the strength and $f_{\rm ext}$ the spatial variation of the ordering field.

These interactions define a coarse-grained model that can be very efficiently studied by computer simulations because (i) the interactions are soft, (ii) the absence of harsh excluded volume interactions allows for a high polymer density and an invariant degree of polymerization that is comparable to experimental values, and (iii) the calculation of the non-bonded interactions via a coarse-grained grid speeds up the calculation of the energy by about two orders of magnitudes in dense systems.[12] The model is simulated by Single-Chain-in-Mean (SCMF) simulations[12,11] where the explicit chain conformations evolve in time via a Smart Monte Carlo algorithm.[17]

For this model, the free energy changes along the two branches depicted in Fig. 1 take the simple, explicit form

$$\frac{\Delta F_1}{nk_BT} = -\int_0^{\Lambda N}\mathrm{d}\lambda N\ \left\langle \frac{1}{N_{\rm cells}}\sum_\mathbf{j} f_{\rm ext}(\mathbf{j})\phi(\mathbf{j})\right\rangle\Bigg|_{\chi_{\rm init}N} \tag{2}$$

$$\frac{\Delta F_2}{nk_BT} = -\int_{\Lambda N}^{0}\mathrm{d}\lambda N\ \left\langle \frac{1}{N_{\rm cell}}\sum_\mathbf{j}\left(f_{\rm ext}(\mathbf{j})\phi(\mathbf{j}) - \phi^2(\mathbf{j})\right)\right\rangle \tag{3}$$

We simulate the systems for different state points along the integration path, measure the integrands of Eqs. (2) and (3), and evaluate the free energy changes by numerically evaluating the integrals. This provides a first estimate for the free energy changes along the path. Subsequently, we employ an expanded ensemble[18] where the incompatibility, χN, and the strength of the external field, λN, vary along the integration path. Within a single simulation run the system will visit different state points along the integration path. To this end, the Monte Carlo algorithm comprises additional moves that alter χN and λN. The expanded partition function takes the form

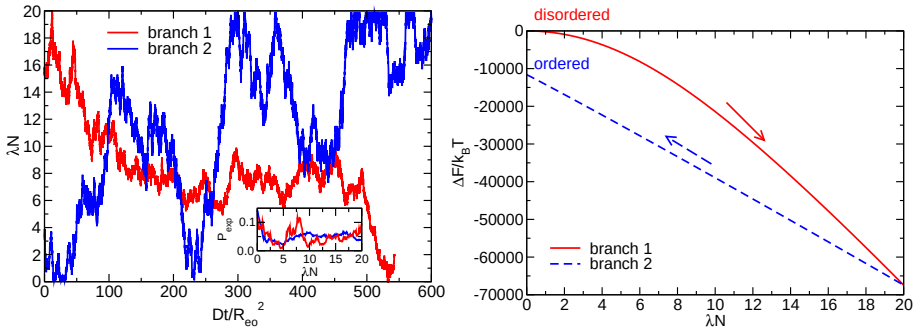


Figure 2. a) Evolution of the ordering field, λN, in the course of the expanded ensemble simulation along both branches. The "time" is measured in units of the Rouse-time of the macromolecules. The inset presents the probability distribution, $P_{\rm exp}(\chi N, \lambda N)$. b) Changes of the free energy along the two branches of the integration path.

$$\mathcal{Z}_{\rm exp} = \sum_{(\chi N, \lambda N)} \exp\left[\frac{w(\chi N, \lambda N)}{k_B T}\right] \int \mathcal{D}[\{\mathbf{R}_i\}] \exp\left[-\frac{\mathcal{H}_{\rm liq} + \mathcal{H}_{\rm ord}(\chi N) + \mathcal{H}_{\rm ext}(\lambda N)}{k_B T}\right] \tag{4}$$

The pre-weighting factors, $w(\chi N, \lambda N)$, facilitate transition between different states, $(\chi N, \lambda N)$, along the integration path. The probability, $P_{\rm exp}$, of finding the system in a state characterized by χN and λN is related to the free energy via $P_{\rm exp}(\chi N, \lambda N) = \frac{1}{\mathcal{Z}} \exp\left[-\frac{F(\chi N, \lambda N) - w(\chi N, \lambda N)}{k_B T}\right]$. The choice, $w(\chi N, \lambda N) \approx F(n, V, T, \chi N, \lambda N)$, for the pre-weighting factors ensures that the different state points are sampled with approximately equal probability. Initial estimates of the pre-weighting factors are obtained by thermodynamic integration and successive umbrella sampling[19] but alternative schemes[20,21] can be envisioned. Note that compared to other re-weighting techniques (utilized e.g., to calculate phase diagrams, interface tensions, or potentials of mean force)[1] the free energy difference along the integration path is large, $\mathcal{O}(10^4 k_B T)$, and a systematic method for obtaining/improving the pre-weighting factors is required.

3 Results

In order to calculate the free energy difference between the disordered state, $\chi N = 0$, and the lamellar ordered structure at $\chi N = 20$, we discretize both branches of the integration path. The variation of the integrands along both branches of the path is completely gradual, indicating the absence of a first-order transition. The absence of a first-order transition is corroborated by Fig. 2a, where we show the evolution of the strength of the ordering field, λN, during the course of the expanded ensemble simulation. The simulation data presented correspond to a single configuration that samples all different external field strengths, λN, of a branch.

The system freely diffuses along the λN-axis and there is no "kinetic barrier" between neighbouring λN-states. This observation demonstrates that the regions of configuration

space associated with neighbouring λN-values overlap. We also observe that the system visits all λN-states with roughly equal probability (see inset). This demonstrates that the pre-weighting factors are very accurate and the error in the free energy difference is of the order of a few k_BT which, in turn, is much smaller than the total free energy difference, $\mathcal{O}(10^4)k_BT$.

The SCMF simulations can be very efficiently implemented on a parallel computer because we simulate an ensemble of independent molecules in fluctuating fields.[11,12] Even for the rather small system size considered in this work, 32 processors can be used. If one takes the absence of a free energy barrier along the thermodynamic integration path for granted, it would be beneficial to subdivide the range of external field strength and use a successive umbrella sampling technique.[19] In this case, the ~ 400 points along the two branches can be divided into several sub-branches (typically, we use 32) that overlap at their boundaries. Each sub-branch can be assigned to a different group of processors. Additionally, we have implemented a parallel-tempering scheme to facilitate the relaxation of the systems by exchanging configurations between different sub-branches. This allows us to efficiently employ $32^2 = 1024$ processors.

In Fig. 2b, the variation of the free energy along the two branches is presented. Since we calculate free energy differences, we arbitrarily set the free energy of the initial, disordered state ($\chi N = 0$) to zero, and we have matched the free energy at the end of the first branch with that of the beginning of the second one. From the data we obtain a free energy difference of $\Delta F/k_BT = 11607(10)$ or $\Delta F/nk_BT = -0.87659(75)$ for a lamellar spacing, $L_o = 1.686R_{eo}$. The error estimate refers to the statistical error of the result but does not include a possible systematic over-estimation of the free energy because of the deviation of the lamellar spacing from its equilibrium value due to the finite size of the simulation cell.

4 Outlook

This new simulation technique permits us to accurately calculate free energy differences in self-assembling soft matter. Potential application encompass the determination of phase diagrams and the identification of stable phases. Applying the scheme to a perfectly ordered system and a system with an interface (e.g., between self-assembled phases with different orientations), one can calculate the free energy of grain boundaries and defects. The calculation of the free energy of a grain boundary is illustrated in Fig. 3. Calculating the free energy difference between a disordered system, $\chi N = 0$, and an ordered system, $\chi N = 20$ with and without grain boundary we obtain the free energy difference $\Delta F = 0.032k_BT$/chain. This value corresponds to an interface free energy, $\frac{\gamma R_{eo}^2}{k_BT\sqrt{N}} = 0.16$. This value is slightly smaller than the value 0.21 predicted by SCF theory.[22] This deviation is due to (i) fluctuations that tend to reduce the free energy and (ii) the differences in the models used in SCF theory and SCMF simulations. While SCF theory assumes zero-ranged interactions, the spatial range of interactions in the SCMF simulations is set by the rather coarse grid spacing, ΔL.

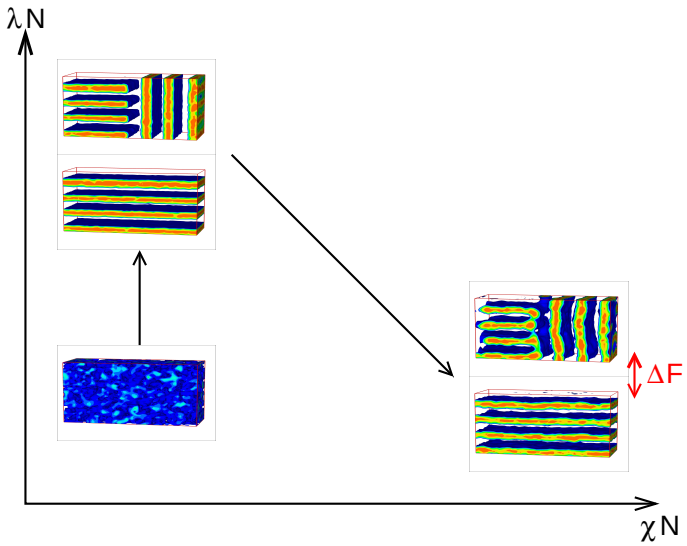


Figure 3. Illustration of the thermodynamic integration path for calculating the free energy difference of a T-junction between two lamellar domains of perpendicular orientation at $\chi N = 20$ and $\bar{N} = 14\,884$. The simulation cell of geometry $4 \times 10.2 \times 6.8 R_{eo}^3$ contains $33\,848$ polymers with $N = 32$ segments.

Acknowledgments

It is a great pleasure to thank M.P. Allen, K. Binder, G.H. Fredrickson, and M.W. Matsen for fruitful discussions. Financial support was provided by the DFG priority program "Nano- and Microfluidics" under grant Mu 1674/3-2 and the Volkswagen foundation. Computing time at the John von Neumann Institute for Computing (NIC) is gratefully acknowledged.

References

1. M. Müller and J. J. de Pablo, *Simulation techniques for calculating free energies*, Lec. Notes Phys., **703**, 67–122, 2006.
2. U. Micka and K. Binder, *Unusual Finite-Size Effects in the Monte-Carlo Simulation of Microphase Formation of Block-Copolymer Melts*, Macromolecular Theory and Simulations, **4**, 419–447, 1995.
3. M. Murat, G. S. Grest, and K. Kremer, *Statics and Dynamics of Symmetric Diblock Copolymers: a Molecular Dynamics Study*, Macromolecules, **32**, 595–609, 1999.
4. C. A. Tyler and D. C. Morse, *Stress in Self-Consistent-Field Theory*, Macromolecules, **36**, 8184–8188, 2003.
5. D. Frenkel and A. J. C. Ladd, *New Monte Carlo Method to Compute the Free Energy of Arbitrary Solids. Application to the FCC and HCP Phases of Hard Spheres*, J. Chem. Phys., **81**, 3188–3193, 1984.
6. S. Y. Sheu, C. Y. Mou, and R. Lovett, *How a solid can be turned into a gas without*

passing through a first-order phase-transformation, Phys. Rev. E, **51**, R3795–R3798, 1995.
7. G. Grochola, *Constrained fluid λ-integration: constructing a reversible thermodynamic path between the solid and liquid state*, J. Chem. Phys., **120**, 2122–2126, 2004.
8. D. M. Eike, J. F. Brennecke, and E. J. Maginn, *Toward a Robust and General Molecular Simulation Method for Computing Solid-Liquid Coexistence*, J. Chem. Phys., **122**, 014115, 2005.
9. M. Müller, K. Katsov, and M. Schick, *Biological and Synthetic Membranes: What Can Be Learned From a Coarse-Grained Description?*, Physics Reports, **434**, 113–176, 2006.
10. G. H. Fredrickson, V. Ganesan, and F. Drolet, *Field-Theoretic Computer Simulation Methods for Polymers and Complex Fluids*, Macromolecules, **35**, 16–39, 2002.
11. M. Müller and G. D. Smith, *Phase Separation in Binary Mixtures containing Polymers: a Quantitative Comparison of Single-Chain-in-Mean-Field Simulations and Computer Simulations of the Corresponding Multichain Systems*, J. Polym. Sci. B: Polymer Physics, **43**, 934–958, 2005.
12. K. Ch. Daoulas and M. Müller, *Single Chain in Mean Field Simulations: Quasi-Instantaneous Field Approximation and Quantitative Comparison With Monte Carlo Simulations*, J. Chem. Phys., **125**, 184904, 2006.
13. N. M. Maurits and J. G. E. M. Fraaije, *Mesoscopic Dynamics of Copolymer Melts: From Density Dynamics To External Potential Dynamics Using Nonlocal Kinetic Coupling*, J. Chem. Phys., **107**, 5879–5889, 1997.
14. E. Reister, M. Müller, and K. Binder, *Spinodal Decomposition in a Binary Polymer Mixture: Dynamic Self-Consistent-Field Theory and Monte Carlo Simulations*, Phys. Rev. E, **64**, 041804, 2001.
15. J. G. E. M. Fraaije, *Dynamic Density-Functional Theory for Microphase Separation Kinetics of Block-Copolymer Melts*, J. Chem. Phys., **99**, 9202–9212, 1993.
16. J. W. Eastwood, R. W. Hockney, and D. N. Lawrence, *P3m3dp - The 3-Dimensional Periodic Particle-Particle-Particle-Mesh Program*, Computer Physics Communications, **19**, 215–261, 1980.
17. P. J. Rossky, J. D. Doll, and H. L. Friedman, *Brownian Dynamics As Smart Monte-Carlo Simulation*, J. Chem. Phys., **69**, 4628–4633, 1978.
18. A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *New approach to Monte-Carlo calculation of the free-energy – Method of expanded ensembles*, J. Chem. Phys., **96**, 1776–1783, 1992.
19. P. Virnau and M. Müller, *Calculation of Free Energy Through Successive Umbrella Sampling*, J. Chem. Phys., **120**, 10925–10930, 2004.
20. B. A. Berg, *Multicanonical recursions*, J. Stat. Phys., **82**, 323, 1996.
21. F. G. Wang and D. P. Landau, *Efficient Multiple-Range Random Walk Algorithm to Calculate the Density of States*, Phys. Rev. Lett., **86**, 2050–2053, 2001.
22. D. Duque, K. Katsov, and M. Schick, *Theory Of T Junctions And Symmetric Tilt Grain Boundaries In Pure And Mixed Polymer Systems*, J. Chem. Phys., **117**, 10315–10320, 2002.

Molecular Dynamics Simulations of PVA / Pt(111) Interfaces

Konstantin B. Tarmyshov and Florian Müller-Plathe

Technische Universität Darmstadt
Petersenstr. 20, 64285 Darmstadt, Germany
E-mail: {k.tarmyshov, f.mueller-plathe}@theo.chemie.tu-darmstadt.de

Molecular dynamics simulations are used to study the structure and dynamics of the interfaces between the platinum (111) [Pt(111)] surface and liquid isopropanol or polyvinyl alcohol (PVA) melt oligomers (10 monomers long). The mass and number atoms distributions were found to be very similar. The small discrepancies between isopropanol and PVA are introduced by the chemical bonds between monomers within PVA chains. The differences in orientational structure of O-H and O-C(CH) bond vectors are significantly stronger in immediate proximity to the Pt(111) surface. The chemical bonds between the monomers are also the main reason for such differences. Further from the surface, the orientational structures of Pt(111)/isopropanol or PVA melt are similar. The PVA chains that are directly adsorbed on the surface protrude into the bulk up to 2.25 nm. There exists a region, where adsorbed and non-adsorbed chains mix with each other and, very likely, form hydrogen bonds. One can expect that PVA melt, unlike lqiuid isopropanol, attaches relatively strongly onto the Pt(111) surface.

1 Introduction

Properties of metal/polymer interfaces are crucial to a large number of industrial applications. For instance, coating and adhesion are important in car production, machinery, food packaging, solar power cells, and many others. Thin films with metal nanoparticles distributed in them attract much attention due to their prospective applicability in electronics, optical materials, and sensors.[1] Different polymers, in particular polyvinyl alcohol, are used for stabilizing of metal nanoparticles.[2–5]

Understanding the behaviour of polymers near metal surfaces can provide us with the ability to control necessary interfacial properties in existing applications. This knowledge can also help us to choose components for new uses more effectively. Available experimental techniques can be used to obtain energetic, electronic, and structural information of molecules adsorbed on metal surfaces.[6–9] These techniques are used to study very thin films of molecules (from monolayer to few nanometers) on the surface. Experimental conditions here are usually limited to low temperatures and ultra-high vacuum, as one wants to perform experiments free of any contaminants and in stabilized conformation.

Along with the experimental research molecular simulations have also been used to study interfaces between metals and other organic or inorganic materials. Energetic, electronic, structural, and dynamic properties of interfaces obtained from these theoretical studies can be used to complement experimental results. Molecular simulations can also probe systems and conditions that are inaccessible to experimental techniques. There have been already many attempts to study interfaces of metals with other organic, including polymers, and inorganic molecules by means of quantum calculations (QC)[10–16] and molecular dynamics (MD).[17–24] However, quantum or quantum-hybrid calculations can be used only for a relatively small number of atoms (up to few thousands). Classical MD simulations

cannot be used unmodified for simulations of metal/organic interfaces, and especially for transition metals. The complications in MD simulations are caused by the complex chemical and electrostatic interactions between metals and organic matter, for which there are no simple physical laws. Metal/organic interfaces cannot, therefore, be easily (and systematically) parameterized and simulated in a way as it is done, for instance, in the case of purely organic systems. One has to introduce and parameterize a specific (state-of-the-art) treatment of the interactions between organic matter and transition metal surface.

Delle Site et al. have already studied bisphenol-A-polycarbonate melt near a Ni(111) surface via coarse-grained molecular dynamics simulations.[25] In their system, only ends of polymer chains could be adsorbed on the metal surface chemically. This specific adsorption was found to influence the structure and dynamics of the interface, which are quite different from the interface where no chemical bonding occurs. Schravendijk et al. studied hydrated amino acids at the nickel metal surface using MD simulations, where chemical binding was fitted to the reference quantum calculations.[26]

In our previous contribution we studied Pt(111)/isopropanol (2-propanol) interface,[27] where isopropanol molecules bind to the platinum surface through a weak chemical bond between the hydroxyl oxygen and one of platinum atoms. The purpose of the current work is to make the next step and to study the properties of the interface between the Pt(111) surface and polyvinyl alcohol (PVA) oligomers of different length and at different temperatures. It is also of interest to see how the properties of the interface change depending on the tacticity of oligomers and regular appearance of hydroxyl group in the chains.

2 Motivation for Usage of High Performance Computers

In order to estimate properties of interfaces between organic polymers and metal surfaces using molecular dynamics simulations, one must build a molecular model that represents the real system of interest. For this, it is necessary to create a large enough system that contains not only the immediate interface, but also the *bulk* region of the polymer above the metal surface (Fig. 1). This must to be done, since one is interested in the properties of the interfacial region that is located between the metal surface and the polymer bulk.

To determine if the simulated interface contains the bulk region, one has to calculate certain properties of the bulk that are then compared to those properties of the interface in regions at different distance to the surface. A match between the properies of the bulk and the chosen region must be found. These properties can be static, dynamic, or both. In our previous studies, in which interfaces between liquid isopropanol or PVA melt and platinum (111) [Pt(111)] surface were studied, we have used the profile of the density as a function of the distance from the surface to confirm the presence of the bulk region.[27,28] Ideally, the choice of properties, which are used to determine the bulk region in the simulated system and, therefore, the size of this system, should depend on the properties of the interface that one wants to investigate. However, in the case of polymeric materials near a solid surface, the relaxation times of the polymer can be extremely long[22,27] and, therefore, can pose very high computational demands.

Interactions in coarse-grained simulations are usually short ranged and, thus, can be parallelized using the distributed-memory model and the Domain Decomposition (DD) approach. In atomistic simulations, long range (electrostatic) interactions may be present, which basically ban the usage of the DD strategy for parallelization. In such situation,

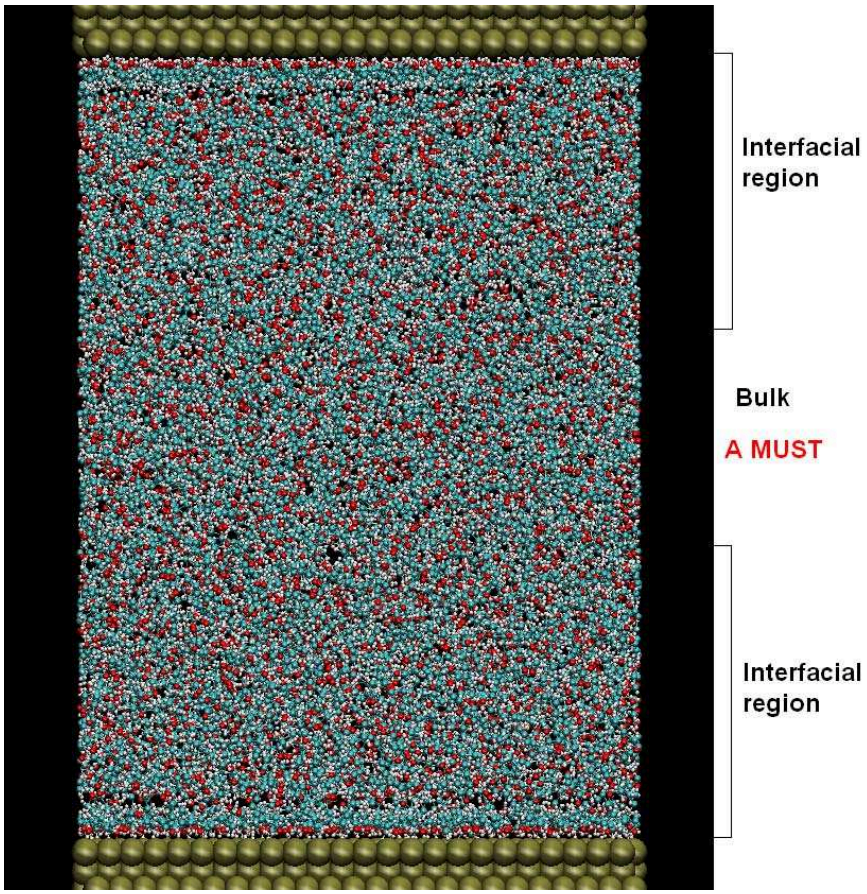


Figure 1. A screenshot of the simulated PVA/Pt(111) interface. The bulk region, which is identified via structural and dynamic properties, **must** be present in any polymer/metal interface system.

the use of shared-memory high performance computers (like ones located at Jülich) for atomistic simulations of interfaces between polymer melts and metal surfaces becomes a very attractive option.

3 Current Work and Results

The interfaces between the Pt(111) surface and liquid isopropanol and melt of PVA oligomers with chain length N=10 have been investigated.[27,28] It was found from the comparison of various structural properties that the structure of the Pt(111)/PVA-melt interface at 400K is similar to that of the Pt(111)/liquid-isopropanol interface at 300K[28]. The distributions of mass density (Fig. 2) and of number density of different atom types (Fig. 3) in PVA melt corespond, generally, to those of liquid isopropanol. The differences in

these distributions can be traced to the chemical bonds between monomers in PVA chains that are absent in the case of isopropanol.

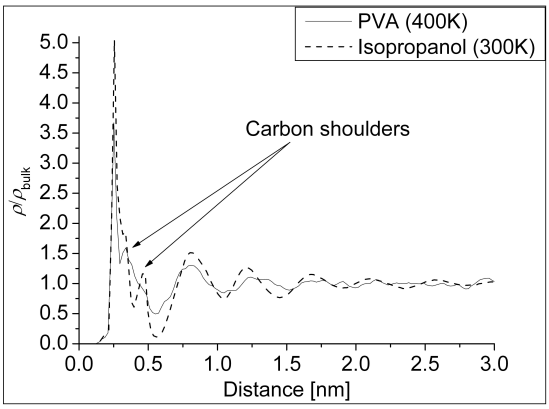


Figure 2. The mass density distribution of PVA melt at 400K and liquid isopropanol at 300K along the surface normal normalized by the corresponding bulk density. The 0 value of the distance axis corresponds to the position of the surface atoms of the platinum slab.

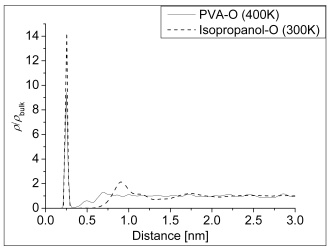


Figure 3a

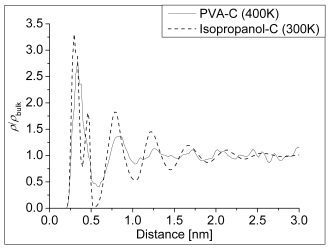


Figure 3b

Figure 3. The number density distribution of hydroxyl oxygen (a) and carbon (b) of PVA melt at 400K and liquid isopropanol at 300K along the surface normal normalized by the corresponding bulk number densities. The 0 value of the distance axis corresponds to the position of the surface atoms of the platinum slab.

This connectivity has substantially stronger impact on the orientational structure. A small change in the distribution of methine carbone atoms (Fig. 4) causes a strong change in the preferred orientation of the O-C(CH) molecular vector (Fig. 5). In the case of isopropanol one can consider that each hydroxyl-methine couple has only methyl groups

as neighbours. For PVA, each of hydroxyl-methine couple (except at the ends of chains) is connected over a methylene bridge to two other hydroxyl-methine couples. It is clear that these hydroxyl-methine couples are larger than methyl groups. Therefore, they introduce additional sterical constraints that result in the sign change of the order parameter S^2 of the O-C(CH) bond vector directly at the surface (Fig. 5).

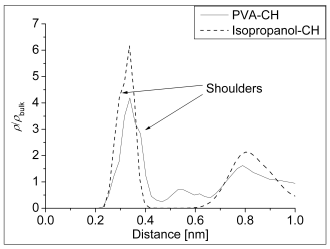


Figure 4a

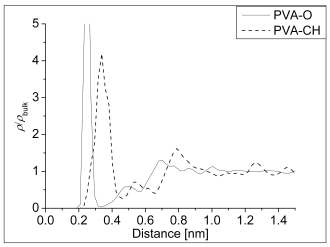


Figure 4b

Figure 4. The number density distribution of the methine carbons for PVA and isopropanol (a), and the hydroxyl oxygens and methine carbons of PVA melt at 400K (b) along the surface normal normalized by the corresponding bulk densities. The 0 value of the distance axis corresponds to the position of the surface atoms of the platinum slab.

The PVA chains adsorbed onto the Pt(111) surface were found to extend into the bulk for up to a few nanometers. There is also an overlap region, where both adsorbed and nonadsorbed chains co-exist and mix with each other (Fig. 6). Hydrogen bonds are very likely formed between adsorbed and nonadsorbed chains in this region. For the Pt(111)/isopropanol interface, interactions between molecules that adsorb directly on the surface (and create hydrophobic methyl brush) and the next layer are weak (Fig. 7). Unlike

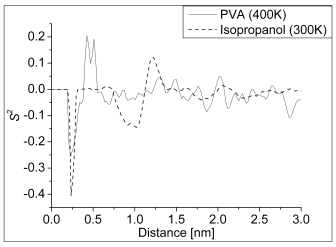


Figure 5a

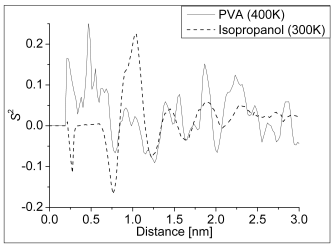
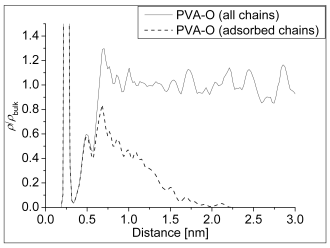


Figure 5b

Figure 5. The order parameter $S^2 = 1/2 < 3cos^2\theta - 1 >$ as a function of distance of molecular vectors O-H (a) and O-C(CH) (b) for melt PVA oligomers at 400K and liquid isopropanol at 300K. The position of the oxygen atom was used to calculate the distance. θ is the angle between the vector being analyzed and the surface normal. $S^2 < 0$ signifies a preferential orientation in the surface plane, $S^2 > 0$ an orientation perpendicular to the surface, and $S^2 = 0$ a random orientation. The 0 value of the distance axis corresponds to the position of the surface atoms of the platinum slab.

in the case of Pt(111)/liquid isopropanol, mixing between adsorbed and non-adsorbed PVA chains will cause stronger attachment of PVA bulk to the adsorbed chains and, therefore, to the Pt(111) surface. However, to quantify this, the non-equilibrium shear molecular dynamics should be carried out.



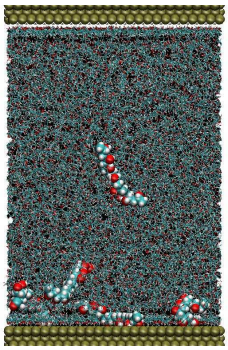


Figure 6a Figure 6b

Figure 6. The number density distribution of hydroxyl oxygen atoms along the surface normal for all chains and only those that have at least one oxygen atom adsorbed onto the Pt(111) surface (PVA melt at 400K) (a). An oxygen atom was considered as adsorbed if its distance was the first minimum of the oxygen density distribution (< 0.32 nm). Both distributions are normalized by the oxygen bulk density in melt of PVA oligomers at 400K. The 0 value of the distance axis corresponds to the position of the surface atoms of the platinum slab.
A snapshot of PVA chains adsorbed on the surface and in the mixing region above the surface (b).

4 Future Work

Currently the bulk properties of PVA oligomers with chain length N=2,3,5, and 7 are being investigated. These studies can be considered as the preliminary step towards the investigation of the properties of interfaces between the Pt(111) surface and mono- and bi-dispersed PVA melts.[29] The obtained properties of the bulk will then be used to build molecular systems with the required bulk region present. These systems will then be used to study the properties of PVA/Pt(111) interfaces and the influence of chain length and mixture composition on these properties.

Acknowledgments

The authors thank Volker C. Weiß for helpful scientific discussions. The authors are thankful to the Hochschulrechenzentrum der Technische Universität Darmstadt for the granted time on the IBM Regatta p575 machines (HHLR) located at the TU Darmstadt, Darmstadt, Germany and also to the John von Neumann-Institut für Computing (NIC) for the granted time on IBM Regatta p690+ machine (JUMP).

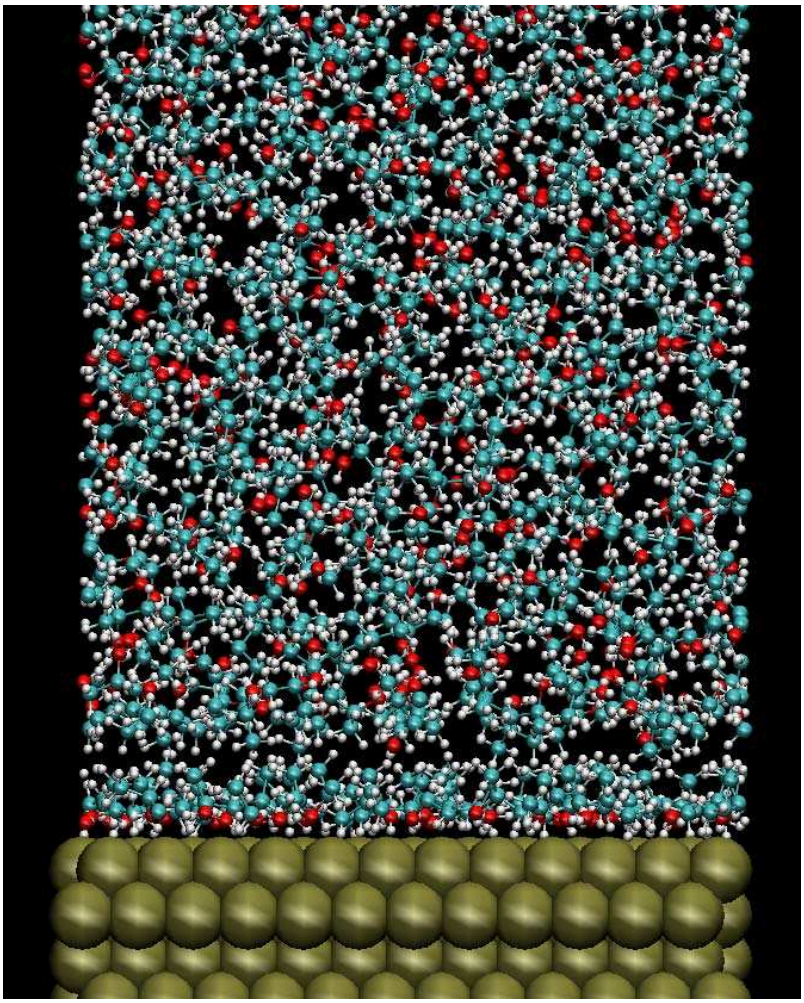


Figure 7. A snapshot of liquid isopropanol adsorbed on the surface. The molecules that are adsorbed directly on the surface create a hydrophobic brush of methyl groups.

References

1. Shipway, A. N.; Katz, E.; Willner, I. ChemPhysChem **1**, 18, 2000.
2. Toshima, N.; Yonezawa, T. New J. Chem. **22**, 1179, 1998.
3. Hirai, H.; Nakao, Y.; Toshima, N. Chem. Lett. , 905, 1976.
4. Hirai, H.; Nakao, Y.; Toshima, N. Chem. Lett. , 545, 1978.

5. Lu, P.; Teranishi, T.; Asakura, K.; Miyake, M.; Toshima, N. J. Phys. Chem. B **103**, 9673, 1999.
6. Temirov, R.; Soubatch, S.; Luican, A.; Tautz, F. S. Nature **444**, 350, 2006.
7. Haq, S.; Harnett, J.; Hodgson, A. Surf. Sci. **505**, 171, 2002.
8. Glebov, A.; Graham, A. P.; Menzel, A.; Toennies, J. P. J. Chem. Phys. **106**, 9382, 1997.
9. Panja, C.; Saliba, N.; Koel, B. E. Surf. Sci. **395**, 248, 1998.
10. Delle Site, L.; Alavi, A.; Abrams, C. F. Phys. Rev. B , 67, 2003.
11. Delle Site, L.; Sebastiani, D. Phys. Rev. B , 79, 2004.
12. Jacob, T.; Goddard, W. A. J. Am. Chem. Soc. **126**, 9360, 2004.
13. Jacob, T.; Muller, R. P.; Goddard, W. A. J. Phys. Chem. B **107**, 9465, 2003.
14. Meng, S.; Wang, E. G.; Gao, S. W. J. Chem. Phys. **119**, 7617, 2003.
15. Meng, S.; Wang, E. G.; Gao, S. W. Phys. Rev. B , 69, 2004.
16. Mittendorfer, F.; Hafner, J. Surf. Sci. **472**, 133, 2001.
17. Lee, S. H.; Kim, H. S. Bull. Korean Chem. Soc. **17**, 700, 1996.
18. Kohlmeyer, A.; Witschel, W.; Spohr, E. Chem. Phys. **213**, 211, 1996.
19. Raghavan, K.; Foster, K.; Motakabbir, K.; Berkowitz, M. J. Chem. Phys. **94**, 2110, 1991.
20. Zhu, S. B.; Philpott, M. R. J. Chem. Phys. **100**, 6961, 1994.
21. Jhon, Y. I.; Kim, H. G.; Jhon, M. S. J. Colloid Interface Sci. **260**, 9, 2003.
22. Abrams, C. F.; Delle Site, L.; Kremer, K. Phys. Rev. E , 67, 2003.
23. Schravendijk, P.; van der Vegt, N.; Delle Site, L.; Kremer, K. ChemPhysChem **6**, 1866, 2005.
24. Witek, H.; Buch, V. J. Chem. Phys. **110**, 3168, 1999.
25. Delle Site, L.; Abrams, C. F.; Alavi, A.; Kremer, K. Phys. Rev. Lett. , 89, 2002.
26. Schravendijk, P.; Ghiringhelli, L. M.; Delle Site, L.; van der Vegt, N. J. Phys. Chem. C **111**, 2631, 2007.
27. Tarmyshov, K. B.; Müller-Plathe, F. J. Chem. Phys. **126**, 74702, 2007.
28. Tarmyshov, K. B.; Müller-Plathe, F. Soft materials **5**, 135, 2007.
29. Rossinsky, E.; Tarmyshov, K. B.; Müller-Plathe, F., 2007, submitted.

Effective Protein Interactions in a Coarse-Grained Model for Lipid Membranes

Beate West and Friederike Schmid

Fakultät für Physik
Universität Bielefeld
33615 Bielefeld, Germany
E-mail: {west, schmid}@physik.uni-bielefeld.de

A simple coarse-grained model for self-assembling lipid membranes is presented. The "lipids" are represented by short linear spring-bead chains, which self-assemble to membranes due to the presence of a computationally cheap "phantom" solvent environment. These membranes may contain "transmembrane proteins", represented by cylinders with diameters corresponding to the diameter of an α-helix. The system is studied by Monte Carlo simulations at constant pressure using a parallel code with a newly devised domain decomposition scheme. The effective interactions between two proteins are calculated for different lipid-protein interactions and compared with the predictions of an elastic theory.

1 Motivation

Biomembranes play a central role in both the structure and function of all biological cells[8]. They serve as an interface between different areas within a cell. A biomembrane, however, is not only a passive interface, but it plays an active role in the transport of molecules and information from one side of the cell to the other.

The biomembrane consists of a liquid-like bilayer of amphiphile lipids, into which membrane-proteins and other macromolecules are inserted. These proteins, e.g. receptors, enzymes and ion channels, are the biochemically active components. Their functions are very versatile: transport (exchange of material through the membrane), enzyme activity, signal transmission (receptors), cell connection, cell-cell recognition.

Both lipid-mediated interactions between proteins included in the membrane and the influence of such inclusions on lipid bilayers, have been intensively investigated for some time[9, 15, 16]. On the one hand the indirect interaction effects obtained by the lipids contribute significantly to the entire interaction between membrane proteins[3, 11]. On the other hand the direct electrostatic interaction effects in biologically relevant constellations become predominantly shielded by the aqueous environment of the membrane[2, 22].

Proteins affect the lateral structure of the lipid membrane: the number of possible conformations of the lipids in the proximity of proteins is reduced[3]. In multicomponent membranes, some lipid components are enriched[7]. Particularly dramatic effects can occur in two-phase areas – with the presence of sufficiently many proteins, the two phases may mix under certain circumstances, and one observes instead of a phase separation a heterogeneous structure of small domains[17, 18]. A protein-induced phase separation is also possible under certain circumstances[19]. Such phenomena have great practical importance, since the lateral structure of membranes is closely connected to its functionality. Apart from direct protein-protein interactions there are also different factors that can induce indirect interactions[9]. These are in particular:

1. Hydrophobic mismatch. This factor will come into play, if the length of the hydrophobic region of a transmembrane protein and the thickness of the membrane do not fit.
2. Disturbance of the local structure. The lipids in the environment of a protein lose their freedom of translation or conformation.
3. Membrane fluctuations. Proteins or other inclusions affect the fluctuation spectrum of membranes and limit it.

Different kinds of membrane proteins provide different characteristics of biomembranes. Protein interactions are of great importance to the functionality of the membrane. The indirect, lipid mediated interactions contribute significantly to the entire interaction between membrane proteins.

2 Coarse-Grained Bilayer Model

Our model consists of a self-assembled bilayer of lipids in a solvent environment[6]. The lipids are represented by chains of six tail beads of diameter σ_t and one slightly larger head-bead of diameter σ_h.

Beads not connected with each other interact via a truncated and shifted Lennard-Jones potential:

$$V_{\text{LJ, shifed}}(r) = \begin{cases} \epsilon \left(\left(\frac{\sigma}{r}\right)^{12} - 2\left(\frac{\sigma}{r}\right)^{6} + V_{\text{shift}} \right) & \text{, if } r < r_c \\ 0 & \text{, otherwise} \end{cases} \tag{1}$$

The parameter σ is the mean value of the diameters of the interacting beads. Head-head and head-tail interactions are purely repulsive ($r_c = \sigma$) and tail-tail interactions also have an attractive part ($r_c = 2\sigma$). The adjacent beads of the lipid chain are bound to each other by a finite extensible nonlinear elastic potential (FENE potential):

$$V_{\text{FENE}}(r) = -\frac{1}{2}\epsilon(\Delta r_{\text{max}})^2 \log\left(1 - \left(\frac{r - r_0}{\Delta r_{\text{max}}}\right)^2\right) \tag{2}$$

Additionally, chains are given a bending stiffness by a bond-angle potential:

$$V_{BA}(\theta) = \epsilon(1 - \cos(\theta)) \tag{3}$$

The solvent environment is represented by explicit solvent beads[13]. They behave like unbounded head beads, except for not interacting with each other. Figure 1 shows the phase diagram of our model[12].

3 Coarse-Grained Protein Model

The proteins are modelled as cylinders. The diameter of these cylinders correspond to that of an α-helix. The proteins are free to move in the xy-plane. The interaction in the xy-plane is represented by a Lennard-Jones kind of potential:

$$V_{\text{LJ, shifed}}(r) = \begin{cases} \epsilon \left(\left(\frac{\sigma}{r - r_{ir}}\right)^{12} - 2\left(\frac{\sigma}{r - r_{ir}}\right)^{6} + V_{\text{shift}} \right) & \text{, if } r - r_{ir} < r_c \\ 0 & \text{, otherwise} \end{cases} \tag{4}$$

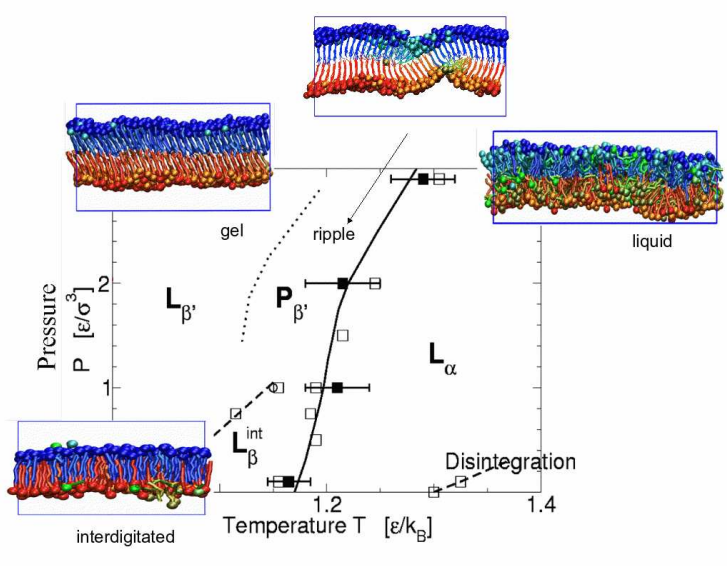


Figure 1. Phase diagram of our lipid model.

The interaction between the proteins and the head and solvent beads is purely repulsive. The interaction with the tail beads is also repulsive. For the tail-protein interaction there is an additional attractive component that depends on the z-distance between the tail bead and the protein.

4 Methods

The system is simulated using Monte Carlo methods at constant pressure and temperature with periodic boundary conditions. The simulation box is allowed to fluctuate during the simulation, i.e., we have additional volume and shape moves besides the moves of the beads.

The programme is parallelised using a geometrical decomposition[20,21]. The idea is to define "active regions", represented by the light blue areas in figure 2. Each processor gets one of these active regions. The distance between the regions is a little bit larger than the maximum interaction range of the beads. This is important to avoid interactions between beads in various active regions. Only the particles inside the active regions will be moved during a Monte Carlo step. Moves out of the active region will be rejected. To make sure that ergodicity is fulfilled, the offset of the active regions, represented by the red arrow in figure 2, is regularly moved randomly.

5 Lipid Bilayer with Two Proteins in the Fluid Phase

We simulate a lipid bilayer with two proteins in the fluid phase L_α for different protein-tail interaction strengths ϵ_{pt}, figure 3. The initial distance of the proteins is chosen such

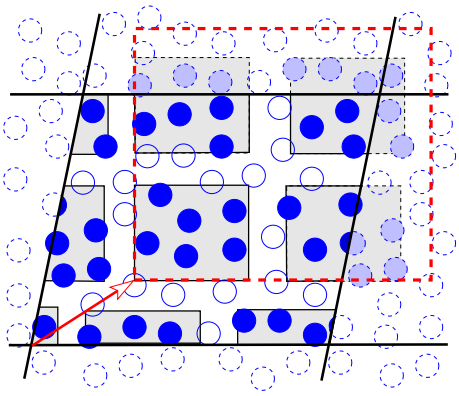


Figure 2. Parallelisation using a geometrical decomposition scheme.

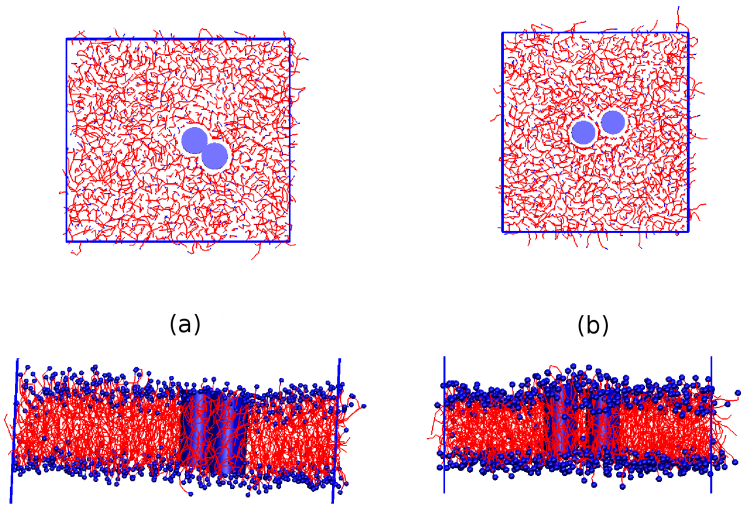


Figure 3. Lipid bilayer with two proteins in the fluid phase L_α: (a) weak interaction between proteins and tails (b) strong interaction between proteins and tails.

that they just touch each other. If the protein-tail interaction is weak, the proteins will stay next to each other and move together during the whole simulation. If the protein-tail interaction is stronger, the proteins will remain separated by a layer of one or two lipids. The membrane clearly is curved around the proteins.

From the pair distribution function $g(r)$ as a function of the protein-protein distance we can calculate the effective pair protein-protein interaction *via*:

$$w(r) = -k_B T \ln g(r) \tag{5}$$

$w(r)$ is the effective pair potential, k_B is the Boltzmann constant and T is the tempera-

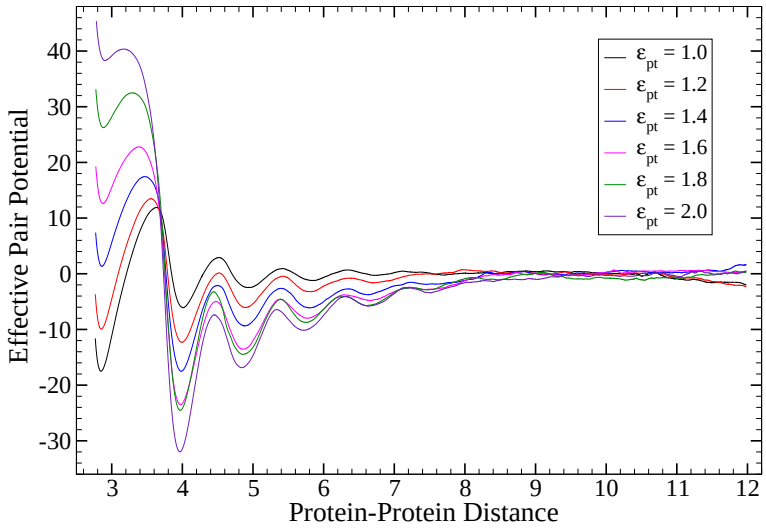


Figure 4. Effective pair potential between two proteins in the fluid phase.

ture. The pair distribution function is obtained by an umbrella sampling procedure as a function of the protein-protein distance. We first make several independent simulations i (preruns), where we contrain the protein-protein distance to stay within a given range $r \in [r_{min,i}, r_{max,i}]$. The windows $[r_{min,i}, r_{max,i}]$ overlap. From these prerun simulations, we deduce an estimate for the distribution of distances $h(r)$ in each window. Then we make a second run in each window, using $1/h(r)$ as a reweighting function to improve the statistics in the valleys. This gives us unnormalized pieces of $g(r)$ in each window. Finally, we put all these pieces together. The result is shown in Figure 4 for different values for the protein-tail interactions strength.

When analysing the effective pair potential different mechanisms seem to be important. On the one hand there is a layering effect. The proteins prefer distances where they are separated by a layer of one, two, ... lipids. On the other hand, there is a smoother interaction contribution, which presumably comes from a hydrophobic mismatch – the configuration snapshots (Figure 3) show that the membrane thickness is enhanced in the vicinity of the proteins. This lowers the curves if the protein-tail interaction is high. The first minimum corresponds to the region where the proteins are in direct contact. Note that the direct interaction between proteins is purely repulsive; however, an attractive depletion interaction is mediated by the solvent and the lipids. The final behaviour of the proteins is determined by the interplay of the different factors. If the tail-protein interaction is low, the effective contact interaction dominates. If the tail-protein interaction is high, the hydrophobic mismatch energy seems to dominate.

To assess the latter more quantitatively, we relate our results with the predictions of the elastic theory of membrane-induced interactions between inclusions by Aranda-Espinoza et al[1]. This theory describes a membrane as a system of two coupled elastic sheets (monolayers), taking into account their area compressibility, bending rigidity, and the individual spontaneous curvature. To extract the elastic parameters in our system, we have determined

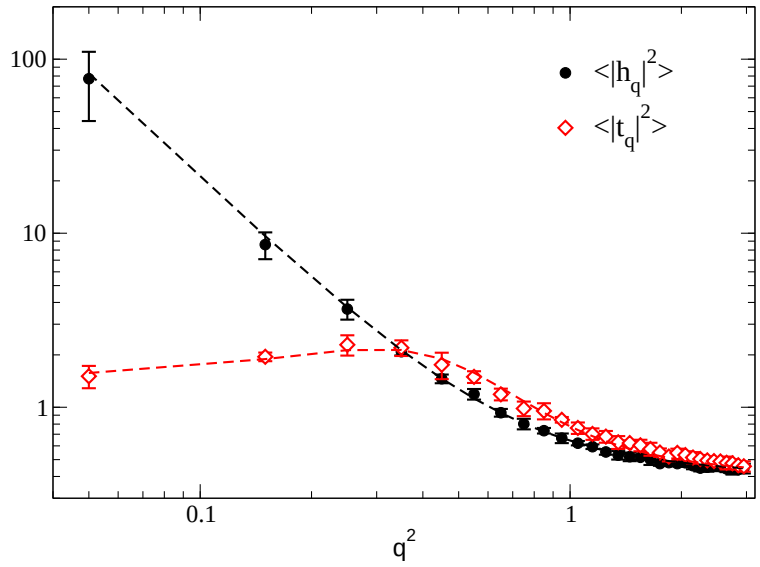


Figure 5. Capillary wave spectrum $< |h_q|^2 >$ (black circles) and thickness fluctuation spectrum $< |t_q|^2 >$ (red diamonds) of pure lipid membrane with fit to the elastic theory[4] (dashed lines). Both curves have been fitted simultaneously.

the fluctuation spectra of membrane position (capillary waves) and membrane thickness in pure membranes and fitted them to the elastic theory of Brannigan and Brown[4]. Figure 5 shows that the fit works very nicely. It provides values for the bending rigidity and the compressibility modulus. The latter is consistent with independent measurements of the lipid area as a function of membrane tension[10]. The remaining elastic parameter, the spontaneous curvature, has been calculated from the first moment of the pressure profile[14].

The elastic theory not only makes predictions for the interaction between two proteins, but also for the distortion of the membrane close to a single protein. Therefore, we have also studied a system with only one protein, and evaluated the thickness profile $u(r)$, *i.e.*, the distance between head beads in the upper and lower layer, as a function of the distance to the protein r. Figure 6 (left) shows that the hydrophobic mismatch seems to play an important role. At strong protein-tail interaction the membrane gets distorted.

The theory captures the distortion, but quantitatively, it does not agree with the simulation data. It predicts strong oscillations, which are not seen in the simulations. We have also evaluated the effective interaction potential (Figure 6, right). Even taking into account that the theory is not designed to capture layering effects, the agreement is still not good.

Hence the simple, straightforward application of the elastic theory fails to explain the findings of our simulation. More refined versions of the theory[5] might provide a remedy. It is also possible that additional factors have to be taken into account. For example, the vicinity of the protein will most likely affect the spontaneous curvature locally.

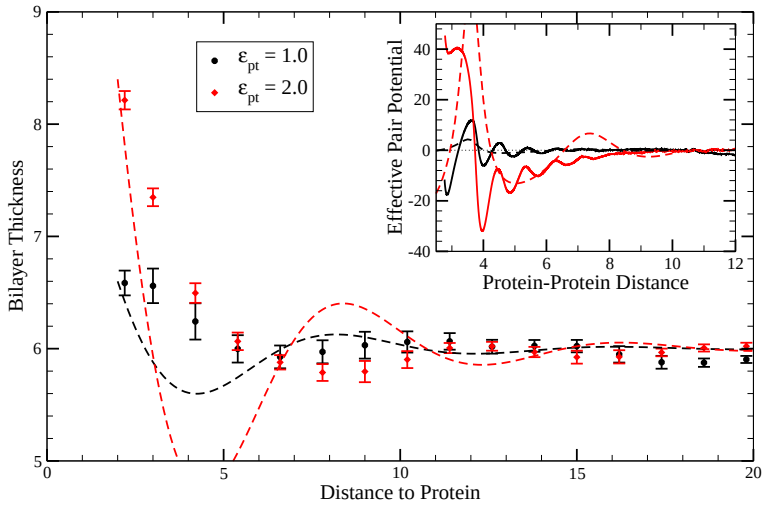


Figure 6. Comparison of the simulation data with the predictions of the elastic theory. Main panel: Thickness of a lipid bilayer depending on the distance to the protein for different protein-lipid interactions as indicated. Points are simulation data, dashed lines correspond to the theory with one fitting parameter, the relative distortion Δ_0 at contact. Inset: Effective protein-protein interactions for the same parameter values. Data are the same as in Figure 4 (solid lines). Dashed lines indicate theoretical prediction for the same values of Δ_0 as in the main panel.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft within the Sonderforschungsbereich 613 “Physik von Einzelmolekülprozessen und molekularer Erkennung in organischen Systemen” and was supported with computing time by the John von Neumann Institute for Computing.

References

1. H. Aranda-Espinoza, A. Berman, N. Dan, P. Pincus, and S. Safran, *Interaction between Inclusions Embedded in Membranes*, Biophysical Journal **71**, 648, 1996.
2. N. Ben-Tal, and B. Honig, *Helix-helix interactions in lipid bilayers*, Biophysical Journal **71**, 3046, 1996.
3. M. Bloom, E. Evans, and O.G. Mouritsen, *Physical properties of the fluid lipid-bilayer component of cell membranes: a perspective*, Quart. Rev. Biophys. **24**, 293, 1991.
4. G. Brannigan and F.L.H. Brown, *A Consistent Model for Thermal Fluctuations an Protein-Induced Deformations in Lipid Bilayers*, Biophysical Journal **90**, 1501, 2006.
5. G. Brannigan and F.L.H. Brown, *Contributions of Gaussian Curvature and Nonconstant lipid Volume to Protein Deformation of Lipid Bilayers*, Biophysical Journal **92**, 864, 2007.
6. D. Düchs, and F. Schmid, *Phase behavior of amphiphilic monolayers: Theory and simulations*, J. Phys: Cond. Matt. **13**, 4853, 2001.

7. F. Dumas, M.M. Sperotto, M.C. Lebrun, J.F. Tocanne, and O.G. Mouritsen, *Molecular sorting of lipids by bacteriorhodopsin in dilauroylphophatidylcholine/distearoylphosphatidylcholine lipid bilayers*, Biophysical Journal **73**, 1940, 1997.
8. R.B. Gennis, *Biomembranes - Molecular Structure and Function*, Springer Verlag, New York, 1989.
9. T. Gil, J.H. Ipsen, O.G. Mouritsen, M.C. Sabra, M.M. Speretto, and M.J. Zuckermann, *Theoretical analysis of protein organization in lipid membranes*, Biochimica and Biophysica Acta **1376**, 245, 1998.
10. J. Neder, private communication.
11. P. Lague, M.M. Zuckermann, and B. Roux, *Protein inclusion in lipid membranes: A theory based on the hypernetted chain integral equation*, Faraday Disc. **111**, 165, 1998.
12. O. Lenz, and F. Schmid, *Structure of Symmetric and Asymmetric Ripple Phases in Lipid Bilayers*, Phys. Rev. Lett. **98**, 058104, 2007.
13. O. Lenz, and F. Schmid, *A simple computer model for liquid lipid bilayers*, J. Mol. Liquids **117(1-3)**, 147, 2005.
14. D.O. Marsh, *Elastic curvature constants of lipid monolayers and bilayers*, Chemistry and Physics of Lipids **144**, 146, 2006.
15. S. May, *Theories on structural perturbations of lipid bilayers*, Curr. Opn. Coll. Interf. Sci. **5**, 244, 2000.
16. O.G. Mouritsen, and M. Bloom, *Models of lipid-protein interactions in membranes*, Quarterly Reviews of Biophysics: Biomol. Struct. **22**, 145, 1993.
17. B. Piknova, D. Marsch, and T.E. Thompson, *Fluorescence quenching and electron spin resonance study of percolation in a two-phase lipid bilayer containing bacteriorhodopsin*, Biophysical Journal **72**, 2660, 1997.
18. V. Schram, and T.E. Thompson, *Influence of the intrinsic membrane protein bacteriorhodopsin on gel-phase domain topology in two-component*, Biophysical Journal **72**, 2217, 1997.
19. P. Sens, and S.A. Safran, *Inclusions induced phase separation in mixed lipid films*, European Physical Journal E **1**, 237, 2000.
20. F. Schmid, D. Düchs, O. Lenz, and B. West, *A generic model for lipid monolayers, bilayers, and membranes*, Computer Physics Communications **177(1-2)**, 168, 2007.
21. A. Uhlherr, S.J. Leak, N.E. Adam, P.E. Nyberg, M. Doxastakis, V.G. Mavrantzas, and D.N. Theodorou, *Large scale atomistic polymer simulations using Monte Carlo methods for parallel vector processors*, Computer Physics Communications **144(1)**, 1, 2002.
22. D. Walther, P. Kuzmin, and E. Donath, *Brownian dynamics simulation of the lateral distribution of charged membrane components*, Eur. Biophys. J. **24**, 125, 1996.

Earth and Environment

Ulrich Hansen

Institut für Geophysik, Westfälische Wilhelms-Universität Münster
Corrensstr. 24, 48149 Münster, Germany
E-mail: hansen@earth.uni-muenster.de

Climate change is clearly the phenomenon at the centre of research efforts in environmental sciences at the moment. We have discovered that the climate is changing and that in fact it seems to change faster than ever. The reasons behind the climatic variations are not clear yet and are still debated. Especially the impact of man on the climate, or on System Earth, in a broader sense, is subject of converse discussions. "The footprint of mankind upon Earth is too deep" is one opinion, as correspondingly formulated by the Nobel laureate Al Gore. Others favour the view, that the variations, as momentarily witnessed, are simply due to natural fluctuations in a complex system. These rather opposing perspectives reflect, that our understanding of System Earth is far from being complete. One particular reason is, that we are affected by processes occurring on an immense spectrum of spatial and temporal scales. Furthermore, we begin to understand that the different geospheres, i.e. hydrosphere, atmosphere and the interior of the Earth are closely coupled. A strong volcano eruption can significantly influence atmospheric conditions on Earth. On the other hand, the presence of an early atmosphere would have influenced the thermal evolution of our planet's interior. Besides changes in climate we have recently experienced a series of events, dramatically influencing human life and habitat. Especially the big Sumatra Earthquake on December 26 th 2006 has given us an idea, of how violent "natural fluctuations" can be. Such an event unhinges the view of an Earth, peaceful, if not brought out of balance by it's inhabitants. Going back through Earth's history reveals several global events which have been termed "Evolutionary Catastrophes". At least twice, 250 and 65 Million years ago, the vast majority of all inhabitants disappeared from the Earth's surface. Up to now it is unclear if these mass extinctions were caused externally, by for example an asteroid impact, or, if the forces originated from deep within the Earth. Seemingly it does not take humans to generate hazards of even global scale. There is in fact little hope that such events can be influenced by any means, hopefully being different from the threads arising from climatic changes. Still, research has tremendously improved our knowledge with respect to the dynamics of our habitat.

Powerful computers have become the most important research tools. While laboratory experiments hardly allow to encompass all scales involved, this becomes more and more feasible in numerical models. Even today's most up-todate models can not resolve all relevant scales. However, together with a further developed understanding of fundamental properties of complex systems, one can hope for models which are able to capture the essential dynamics of Oceans, Atmosphere and Earth's interior. Computers also allow to take a look into the interior of our planet, possibly inaccessible forever. Tomographic methods have been developed, making use of waves generated by earthquakes, much as using X-rays in medical tomography. But the quest goes further. We do not only want to understand

what the structure is - we want to understand what processes are operating and what are the consequences and results of these processes. The behaviour of the material under such extreme conditions is one key ingredient towards an understanding of this processes.

The contribution of Jahn, “Fluids Under Extreme Conditions of Pressure and Temperature and their Role in Geological Processes”, aims at a deeper understanding of geomaterials at high pressure and high temperatures. Atomic structures of geomaterials vary significantly from metallic iron, deep in the Earth’s core, to silicates and oxides in the Earth’s mantle and crust. In the presence of fluids ionic and molecular species do coexist. High pressure fluids can have a strong influence of dynamic processes within the Earth. For example, they can strongly influence the viscosity of the mantle material and locally facilitate creeping flows in the mantle. Another topic concerns the investigation of the silicate melts. Material properties of silicate melts, again especially the viscosity, play a key role for geodynamical processes on all scales, ranging from the thermochemical evolution of a magma chamber and its implications for the formation of igneous deposits to the evolution of the Earth’s mantle on a global scale.

The second contribution by Völker “High-Resolution Modelling of Phytoplankton Distribution and Adaptation” aims at the investigation of possible consequences of a climatic change. By means of a high resolution model the impact of a change in the circulation in the Southern Ocean on the composition of plankton shall be investigated. The main focus will be on variations of the circulation patterns as caused by climatic changes within the next century. The composition of the plankton is closely related to the distribution of carbon between the ocean and the atmosphere. A close investigation of the conditions in the Southern Ocean is only a first step. In further efforts the authors aim to develop a state-of-the-art-model, in order to address similar question for other oceanic areas.

Fluids Under Extreme Conditions of Pressure and Temperature and their Role in Geological Processes

Sandro Jahn

GeoForschungsZentrum Potsdam
Section 4.1 Experimental Geochemistry and Mineral Physics
Telegrafenberg, 14473 Potsdam, Germany
E-mail: jahn@gfz-potsdam.de

The knowledge of phase equilibria and geomaterials properties is essential for understanding the complex physics and chemistry of the Earth. Among the different geomaterials, fluids play a crucial role in many geological processes, but their behaviour at the conditions of the Earth's interior, i.e. at high pressures P and high temperatures T, is not well known. We use *ab initio* molecular dynamics (AIMD) simulations to study the atomic structure and dynamics of such fluids to complement experimental approaches. The simulations do not only help to interpret experimental data but they can predict material behaviour under conditions not accessible to experiment. Due to the substantial computational demand of AIMD simulations, this method has only become feasible with the use of supercomputing facilities.

1 Introduction

Geomaterials, i.e. minerals, melts and volatile fluids, exhibit complex atomic structures. Whereas the Earth's core is supposably composed mainly of metallic iron, the crust and mantle consist mainly of silicates and oxides[1]. Different types of atomic structures and particle interactions pose a great challenge for computer modelling in Geomaterials Science. The substantial increase in computing power during the last decades provides the base for a realistic modelling of geologically relevant materials and processes. A first prerequisite for atomic-scale modelling is an accurate and reliable representation of the particle interactions. Electronic structure methods, e.g. using density functional theory (DFT)[2,3], have proven reliable in this respect since they are transferable between many different types of materials. However, they are computationally very demanding which limits the maximum system size, i.e. number of atoms in the simulation cell, and time scales for dynamic simulations. With the availability of supercomputing facilities, *ab initio* molecular dynamics (AIMD) simulations have become feasible. This provides access to chemically complex systems that are not well described by classical interaction potentials.

We are interested in properties of geomaterials and geological processes at conditions of the Earth's interior, i.e. at high pressure and high temperature. Generally, geologic formations are complex composites of different crystalline (minerals) and disordered (melts, glasses, fluids) materials. Their properties are ultimately determined by the bulk properties of the individual constituents and by the interfacial properties at phase boundaries. Both aspects are subject of our research. Especially for melts and fluids, experimental studies under these extreme conditions are very difficult, and important data, e.g. on the atomic structure and dynamics, are still missing. Molecular modelling is a powerful complementary approach, by which structural and physical properties may be studied simultaneously.

Melts and fluids are especially important for processes in the Earth's crust and upper mantle, which is illustrated in figure 1. When an oceanic plate is subducted beneath the

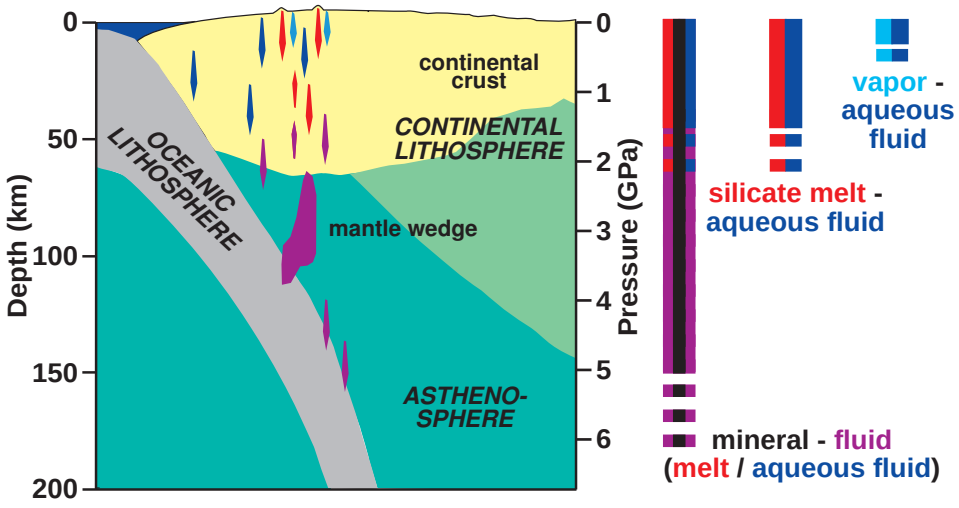


Figure 1. Schematic cross-section through the subduction zone of the Central Andes. Fluids are released at depth due to mineral phase transformations, in particular dehydration reactions, at high pressure and high temperature. The ascending fluids interact with the rocks of the mantle wedge and the continental crust, which leads e.g. to magma formation, volcanism and hydrothermal ore deposit formation.

continental crust, P and T increase with depth. As a result, water and other volatiles are released when low $P-T$ minerals become instable and transform into high $P-T$ phases. The ascending volatiles then cause a partial melting of the mantle wedge or the continental crust. Eventually, fluids and magmas reach the surface again, e.g. during volcanic eruptions. The most important water-bearing liquids of this process contain substantial amounts of silicate and/or electrolyte components. Due to the much lower viscosity, the latter type of fluid is more easily accessible by AIMD. Recently, we have studied the speciation of ions and molecules in aqueous fluid at high pressure and high temperature using supercomputing facilities at NIC Jülich and GFZ Potsdam. The results of these studies will be used to develop a better understanding of the role of fluids on geochemical processes and physical properties of geomaterials.

2 Simulation Method and Usage of HPC Facilities

For a reasonable representation of melt or fluid structure, the simulation cells have to contain at least a few hundred non-identical atoms. Even at high temperature where equilibration is relatively fast, simulation times of at least a few picoseconds are required to obtain converged averages of structural properties. Long simulation times are especially needed when properties of species with low concentration are studied, as for instance ions in solution. The electronic structure calculations are based on density functional theory (DFT) using planewave basis sets[4] as implemented in the CPMD code[5]. The Car-Parrinello MD simulations are performed with a fictitious electron mass of 600 a.u.[6], time steps of 0.12 fs and the BLYP exchange-correlation functional[7,8]. Periodic boundary conditions are applied to represent a bulk system. The CPMD code is used by a number of groups on the

NIC supercomputers JUMP and JUBL and shows good scaling behaviour up to at least 512 processors for the type of calculation that is performed within this project.

In the next section we will present results of two case studies with different system size. The first set of simulations is performed on systems with 64 water molecules and one LiF (or LiOH). Hence the total number of particles is 194 (195), respectively. A typical CPMD calculation with 10000 time steps (corresponding to a time of 1 ps) takes about 19 hours on 32 processors on JUMP. After equilibration, a minimum of five to ten picoseconds are needed to get reasonable averaging, i.e. about one week of continuous computing time for one state point. Using 512 processors on JUBL, 10000 time steps run in about two hours, which gives a speed-up of about 10 although the code runs slightly slower on JUBL than on JUMP using the same number of processors. Comparing the time needed for a single CPMD time step on JUBL, the code scales very well (see table 1).

np	real time [s]	CPU seconds
32	7.30	234
128	2.11	270
512	0.72	369

Table 1. Scaling behaviour with number of processors (np) on JUBL. The times given refer to one CPMD time step of the 194 atom simulation cell (64 H_2O + 1 LiF) at ambient conditions.

In the second study, a considerably larger number of particles is used. The corresponding simulation cell contains 120 H_2O and 4 $MgSO_4$ molecules, which gives a total of 384 atoms. A single CPMD time step takes between 4.65 s and 6.26 s of real time on JUBL using 512 processors, depending on the specific P/T conditions.

3 Atomic Structure of High Pressure Fluids

3.1 Li^+ Coordination Change in Aqueous Fluid with Pressure

A key parameter for a quantitative understanding of fluid properties and their influence on geological processes is the structural arrangement of particles in the fluid. This includes the structure of the majority species or solvent but also the complexation of minority species or solutes. Simple model systems for lithospheric fluids are dissolved salts in aqueous solutions. These systems are quite well known at ambient conditions but not at high P and T. We have used AIMD simulations to study the hydration of a Li cation in H_2O fluid in the NVT ensemble[9]. All simulations are run at constant temperature ($T = 1000\ K$). The pressure is adjusted by changing the volume (density) of the cubic simulation cell according to the equation of state of pure H_2O fluid. As already mentioned above, the simulation cells contain 64 water molecules and one LiF (or LiOH).

The calculations show that structural changes in the fluid influence the first hydration shell of the Li ion. With increasing pressure, the average number of H_2O molecules in the first hydration layer increases considerably from of about three at the lowest simulation pressure (0.3 GPa) to about five at 6 GPa. The distribution of coordination numbers at these two conditions is shown in figure 2. Since there is only a single Li ion in the simulation cell it is apparent that the hydration shell is not a static structural feature but it is

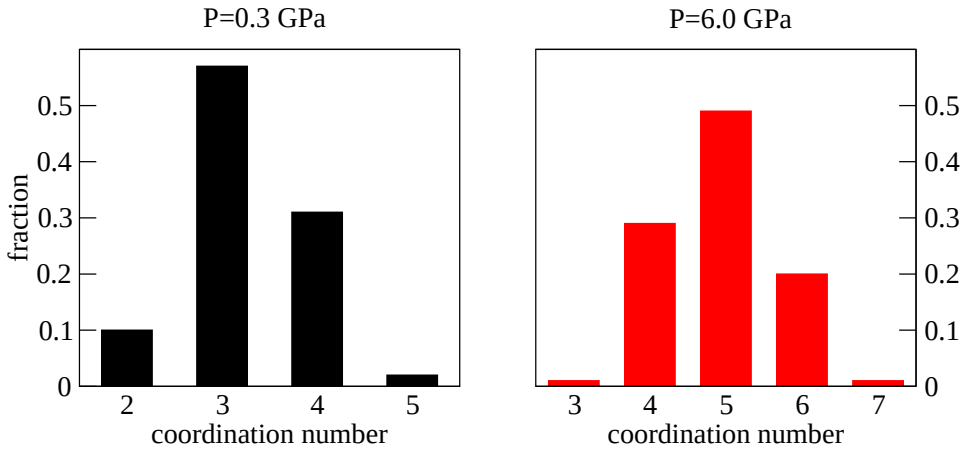


Figure 2. Distribution of the number of anions in the nearest neighbour coordination shell of the Li ion at two different pressures.

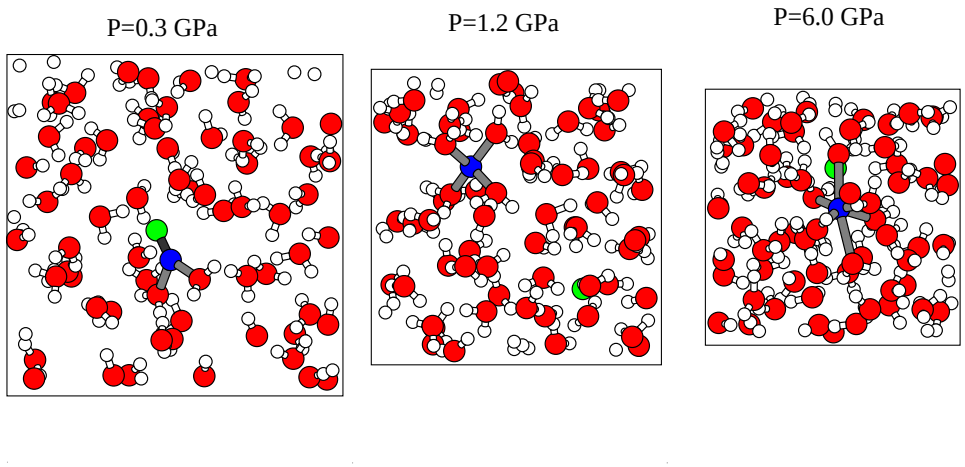


Figure 3. Snapshots from the AIMD simulations demonstrating the increase of Li coordination with increasing pressure. Grey bonds are shown between the Li and the anions of the nearest neighbour shell. Color code: Li-blue, O-red, H-white and F-green.

subject to fast dynamic changes even on the short time scale of AIMD. In figure 3 a few snapshots from the AIMD simulations at different pressures are shown. The compression of the simulation cell with increasing pressure and the increase in Li coordination is readily identifiable. An interesting observation is the tendency of association of LiF or LiOH at low pressures whereas the neutral complexes dissociate at high pressures, which is in qualitative agreement with experiment.

The results of these studies may have impact on the understanding of the geochemical cycles of light elements. For example, the Li^+ coordination change in the fluid has an

effect on the isotopic fractionation between Li-bearing minerals and fluid at high pressure. Experimental studies suggest that different local coordinations in the mineral phases lead to different trends in the fractionation behaviour[10] but so far the pressure induced structural changes are not considered in the geochemical models.

3.2 Speciation of Dissolved $MgSO_4$ in Aqueous Fluid

While the previous study was focused on the hydration of a single ion, real fluids often contain larger ionic or molecular species. Moreover, thermodynamic modelling requires a representative sampling of configurational space, which may be difficult to obtain using single ions. However, the modelling of molecular complexation or polymerization requires much larger simulation cells and hence computing power. In a pilot study we are currently investigating complexation of $MgSO_4$ dissolved in aqueous fluid. This study is in close collaboration with spectroscopists and involves the dissociation and association process of molecular species as a function of pressure and temperature and their relation to the liquid-liquid transition in H_2O fluid.

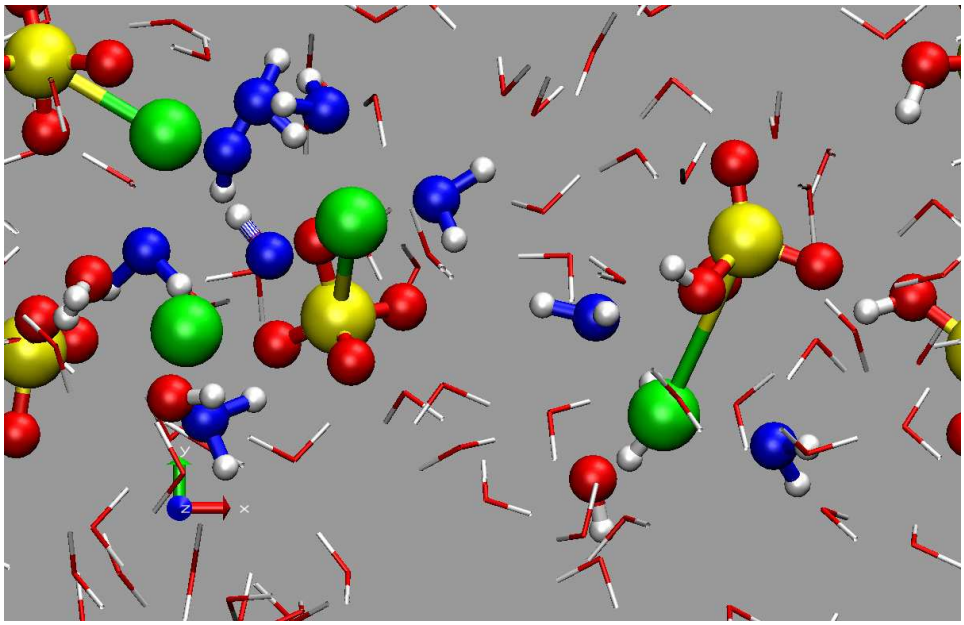


Figure 4. Snapshot from an AIMD simulation of magnesium sulfate in aqueous solution at $T = 1000$ K and $P = 0.5$ GPa. Color code: S-yellow, O-red and blue, H-white, Mg-green. H_2O or OH^- molecules in the first Mg coordination shell are shown as balls and sticks (blue O if they are also in the first hydration shell of the SO_4^{2-} molecule). For associated $MgSO_4$ a bond between S and Mg is drawn.

The speciation of sulfates in water at ambient conditions has been subject to a number of experimental studies[11–14] and quantum chemical calculations of molecular clusters[11,15,16]. The results of the simulations have been used e.g. to identify peaks in the

Raman spectrum with molecular species in the solution[11]. It is not evident how speciation changes at high P/T. It is expected that at high temperature and low pressure ionic species tend to associate[12] and perhaps to polymerize whereas the high pressure behaviour is almost completely unknown. Figure 4 shows a snapshot of an AIMD simulation at $T = 1000$ K and $P = 0.5$ GPa. Three of the four $MgSO_4$ molecules are associated, which is consistent with the expected behaviour. In the simulation we observe some proton jumps from H_2O to SO_4^{2-} which leads to the formation of species containing $[MgOH]^+$ and $[SO_3OH]^-$. These species do not seem to occur in appreciable quantity at ambient conditions. Thus further experimental and simulation studies are needed to understand this observation.

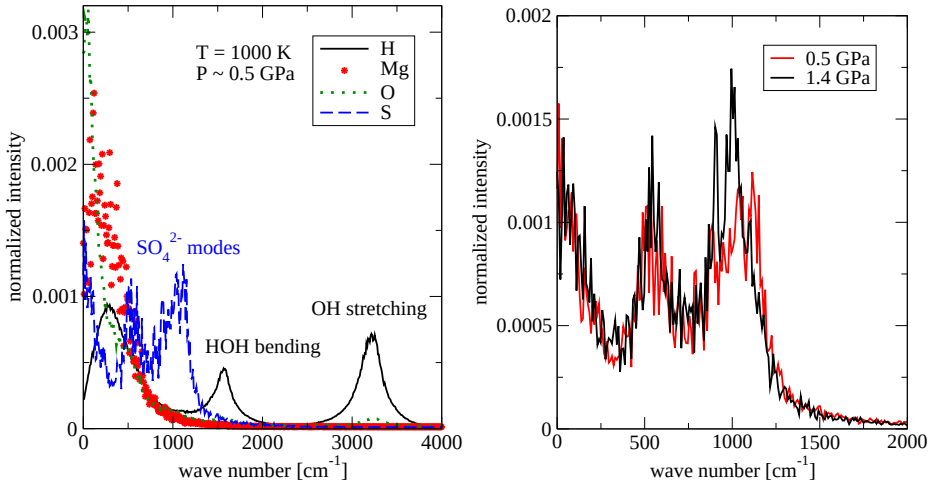


Figure 5. Power spectra of the velocity autocorrelation function representing the vibrational states of the individual species, from AIMD simulations of magnesium sulfate in aqueous solution at $T = 1000$ K. Left: Partial spectra for $P = 0.5$ GPa. Right: Partial spectrum of sulphur at two different pressures.

In order to compare with experimental data, the vibrational spectrum is calculated. Here we present the power spectra of the velocity autocorrelation function

$$z^X(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dt \frac{1}{3}\left\langle \mathbf{v}_1^X(0)\mathbf{v}_1^X(t)\right\rangle \exp(-i\omega t) \qquad (1)$$

where $\mathbf{v}_1^X(t)$ is the velocity of a single atom of species X at time t and ω is the angular frequency. The power spectra yield the characteristic frequencies of the individual particles. In figure 5, the H spectrum shows peaks at wave numbers typical for OH stretching and HOH bending modes. Due to the small number of Mg and S atoms, their corresponding spectra are still quite noisy. However, comparing the S spectra at different pressures, clear differences are observed around 1000 cm^{-1}, which can be related to changes in the speciation. The shift of the spectral peak to lower wave numbers with increasing pressure indicates a tendency of the $MgSO_4$ molecules to dissociate, which is in qualitative agreement with expectations from experiment. A more quantitative analysis of the speciation requires still longer simulation runs, which are currently being performed.

4 Conclusions and Perspectives

The exploratory studies presented here have convinced us that AIMD simulations have great potential to study fluids at geologically relevant conditions and to reveal structural information that is difficult or impossible to access experimentally. These studies are needed not only for thermodynamic modelling of geological systems but they are a prerequisite for understanding the complex processes of fluid-rock interaction. The computational approach to Geomaterials research benefits greatly from HPC facilities such as NIC.

While the accuracy of DFT based simulations is well known for crystalline phases up to extreme conditions of P and T, the respective comparison between model and experiment is much more challenging for fluids and melts. This is partly due to the lack of high quality experimental data at high P and T but also due to the limited length and time scales of the simulations. It is therefore highly desirable to push the limits of both simulation and experiment to allow for maximum overlap and to make best use of the complementary nature of the two approaches.

Acknowledgments

This work has been done in collaboration with experimental geochemists at GFZ Potsdam. I am especially thankful to Christian Schmidt and Bernd Wunder for interesting discussions and useful comments. We are grateful to the *John von Neumann Institute for Computing* (NIC) Jülich for the allocation of CPU time on the supercomputers JUMP and JUBL. We have also used computing resources at GeoForschungsZentrum Potsdam.

References

1. D. L. Anderson, *New theory of the Earth*, Cambridge University Press, 2007.
2. P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev., **136**, B864–B871, 1964.
3. W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev., **140**, A1133–A1138, 1965.
4. R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press, 1989.
5. D. Marx and J. Hutter, *Ab initio molecular dynamics: Theory and implementation*, in: Modern Methods and Algorithms of Quantum Chemistry, Forschungszentrum Jülich, NIC Series, J. Grotendorst, (Ed.), vol. 1, pp. 301–449, 2000.
6. R. Car and M. Parrinello, *Unified approach for molecular dynamics and density functional theory*, Phys. Rev. Lett., **55**, 2471–2474, 1985.
7. A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, **38**, 3098 – 3100, 1988.
8. C. Lee, W. Yang, and R. C. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, **37**, 785–789, 1988.
9. B. Wunder, S. Jahn, and A. Meixner, *Ab initio molecular dynamics study of Li speciation in aqueous fluid at high pressure*, Geochim. Cosmochim. Acta, **71**, A1128, 2007.

10. B. Wunder, A. Meixner, R. L. Romer, and W. Heinrich, *Temperature-dependent isotopic fractionation of lithium between clinopyroxene and high-pressure hydrous fluids*, Contrib. Mineral. Petrol., **151**, 112–120, 2006.
11. C. C. Pye and W. W. Rudolph, *An ab initio and Raman investigation of sulfate ion hydration*, J. Phys. Chem. A, **105**, 905–912, 2001.
12. W. W. Rudolph, G. Irmer, and G. T. Hefter, *Raman spectroscopic investigation of speciation in $MgSO_4(aq)$*, Phys. Chem. Chem. Phys., **5**, 5253–5261, 2003.
13. D. Watanabe and H. Hamaguchi, *Ion association dynamics in aqueous solutions of sulfate salts as studied by Raman band shape analysis*, J. Chem. Phys., **123**, 034508, 2005.
14. S. Schrödle, W. W. Rudolph, G. Hefter, and R. Buchner, *Ion association and hydration in 3:2 electrolyte solutions by dielectric spectroscopy: Aluminum sulfate*, Geochim. Cosmochim. Acta, **71**, 5287–5300, 2007.
15. X. Zhang, Y. Zhang, and Q. Li, *Ab initio studies on the chain of contact ion pairs of magnesium sulfate in supersaturated state of aqueous solution*, J. Mol. Struc. (Theochem), **594**, 19–30, 2002.
16. B. Gao and Z. Liu, *A first-principles study on the solvation and structure of $SO_4^{2-}(H_2O)_n$, n=6-12*, J. Chem. Phys., **121**, 8299–8306, 2004.

High-Resolution Modelling of Phytoplankton Distribution and Adaptation

Martin Losch, Michael Schröder, Sönke Hohn, and Christoph Völker

Alfred-Wegener-Institut für Polar- und Meeresforschung
Am Handelshafen 12, 27575 Bremerhaven, Germany
E-mail: Christoph.Voelker@awi.de

A state-of-the-art eddy-resolving ocean general circulation model has been coupled to a newly developed ecosystem and biogeochemical model that attempts to describe the physiology of phytoplankton cells, their adaptation to changing external conditions and the separate cycling of several phytoplankton nutrients (C, N, Si, Fe). In its initial state the advection of a large number of tracers in the model caused an increase in MPI-exchange costs that inhibited a good scaling behaviour with increasing number of processors used. This has been overcome by bundling the MPI exchange for all tracers. First results from the model show a much improved representation of small-scale biological features, e.g. along the axis of the Antarctic Circumpolar Current, a decoupling of the C and N cycles in nitrogen-limited subpolar gyres, and an increased Si:N ratio in iron-limited regions.

1 Introduction

The marine biology, especially the phytoplankton (suspended microscopic algae and photosynthesizing bacteria), is tightly involved in the geochemical cycling of many elements, especially of carbon, but also compound substances that may have climatic impacts, such as dimethylsulfide or organohalogens. The ocean contains approximately 50 times as much carbon as the atmosphere, and over time-scales longer than a few decades the CO_2 partial pressure of the atmosphere is tightly coupled to the distribution of carbon within the oceans. An important process in regulating this distribution is the so-called 'biological pump', the fixation of carbon through photo-synthesis close to the ocean surface, and the subsequent sinking and decomposition of parts of the produced biological material in the deep ocean. Changes in the growth, mortality and sinking of phytoplankton therefore have the potential to act as a strong feedback on the change in climate that is induced by the anthropogenic accumulation of greenhouse gases in the atmosphere[1–3].

The Southern Ocean is a place of extremely strong exchange fluxes between ocean and atmosphere and plays an important role also for the global carbon cycle[4]. At the same time, it is also a place where strong changes in the ocean circulation are expected under global warming, e.g. a southward shift and strengthening of the Antarctic Circumpolar Current (ACC)[5], a reduction in sea ice cover[6] and a warming of the water column[7].

A good understanding of the feedbacks in this system can only be gained by the use of computer models that allow to predict the physical background that the biology operates in, i.e. the ocean circulation and other external factors that influence phytoplankton growth (nutrients, light), but also the internal regulations within the planktonic ecosystem. Owing to their often large number of advected tracers (and therefore memory requirements) and the longer integration times required, coupled physical-biogeochemical models of the ocean are still often relatively coarsely resolved (on the order of 100-200 km spatial resolution in the horizontal), i.e. they are unable to represent the action of the energetic mesoscale

eddy field that has lengths scales on the order of the Rossby radius of deformation (10-40 km). The mesoscale eddies have been shown to exert a strong influence of the mean circulation, especially in the Southern Ocean[8]. Model predictions of atmosphere-ocean carbon fluxes differ most between different models in the Southern Ocean, owing to a poor representation of the physics there in coarse-resolution models[4]. It is therefore desirable to run coupled physical-biogeochemical ocean models at the state-of-the-art spatial resolution that is used in purely physical modelling.

There are also new challenges for the biogeochemical and ecosystem part of the coupled models. Marine ecosystem models have traditionally worked on the assumptions of a single limiting nutrient, usually nitrogen, for phytoplankton growth, and a constant carbon:nitrogen ratio in biomass[9,10]. However, over the last decade it has become clearer that (i) different phytoplankton functional groups (diatoms, coccolithophorids) have different effects on the carbon flux through the system, (ii) that other nutrients, such as iron and silicate are important regulators of ecosystem structure and functioning, and (iii) that there is some plasticity in phytoplankton physiology, allowing the cells to adapt to different light/nutrient conditions and resulting in variable elemental stoichiometry. We have developed an ecosystem model that focuses mainly on the two last challenges, by implementing a model for diatoms with variable C:N:Si:chlorophyll ratio that depends on the availability of the nutrients N, Si, and Fe. Here we show some results from implementing this model into both a coarser-resolution and a high-resolution global circulation model that have been run on the IBM system *JUMP* at NIC.

2 The Model

2.1 Biogeochemical Model

The ecosystem model is based on the physiological model for the adaptation of algal cells to light, nutrient and temperature by Geider et al.[11] that describes the variation of the cells C:N:chlorophyll ratios with varying conditions assuming balanced growth. This model has been somewhat extended and implemented into a full marine ecosystem model, which describes also the mortality of phytoplankton from zooplankton grazing, the formation of detritus (also through the excretion and aggregation of transparent exopolymeric particles), and other processes by Schartau et al[12].

We have extended this Regulated Ecosystem Model (REcoM)[12] to a model for diatom growth that also includes limitations by silica and iron availability. The parameterization for the uptake of silica and the dependence of growth on the silica quota is based on physiological observations. It results in increased Si:N ratios in the cells under iron-limiting conditions without prescribing this effect explicitly (Hohn et al., in preparation).

2.2 Physical Model and CS510 Configuration

The simulation of the ocean circulation is done using the Massachusetts Institute of Technology general circulation model (MITgcm)[13,14] which solves the time-dependent, Boussinesq-approximated Navier-Stokes equations with or without hydrostatic approximation, and conservation equations for salinity and energy (in the form of an equation for potential temperature). Turbulent diffusion and viscosity are parameterized following the

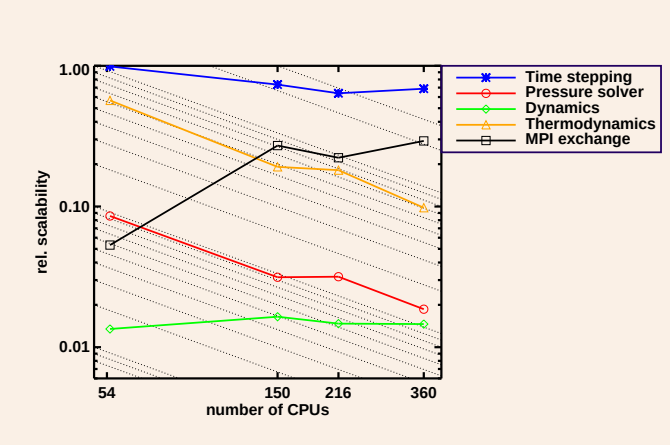


Figure 1. Relative execution times for certain/different parts of the MITgcm model. Here the biogeochemistry module (including 16 tracers) has been switched on. The diagonal lines show an ideal scaling behaviour (computing time inversely proportional to the number of CPUs).

scale-dependent Smagorinsky approach. For the vertical mixing in the surface layer, the KPP model by Large et al.[15] is used.

The equations are discretized using a finite-volume method on an orthogonal grid. To avoid the singularity of a latitude-longitude grid at the poles and to ensure an approximately homogeneous spatial resolution, we have used the 'cubed sphere' grid configuration provided my the MITgcm, which is obtained by mapping a regular grid on the surface of a cube onto a sphere. The high-resolution runs were performed with a horizontal resolution of 510×510 grid points on each of the 6 sides of the cube.

3 Scalability

The MITgcm model has been chosen among other reasons, because of its known good multi-processor scalability over a large range of processor numbers and computer architectures. However, contrary to our previous experiences, the coupled system MITgcm + REcoM shows a bad scaling behaviour on the IBM system *JUMP* with increasing number of processors used. The wall-clock time only weakly decreases with increasing processor number, and even slightly increases in the simulation with 360 processors, compared to a 216-processor configuration. An analysis of the execution times for individual parts of the program (Figure 1) shows the reason for this scaling behaviour. While the main parts of the model (dynamical, thermodynamical and biogeochemical) scale approximately linearly, the required computing time for MPI interprocessor exchange increases with increasing number of processors used. The fact that the execution time for the pressure solver also does not decrease linearly with processor number does not influence the overall execution time significantly, as it is only a small fraction of the total.

A comparison to test without the biogeochemical module (Figure 2) shows that it is the MPI-communication for the 16 passive tracers that is responsible for the bad scaling

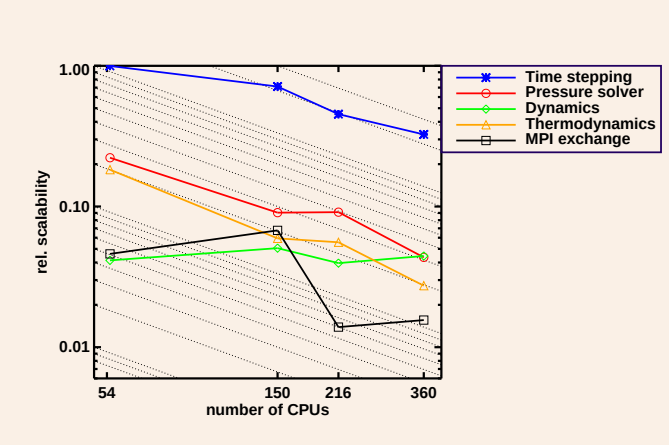


Figure 2. Same as Figure 1, but the biogeochemistry module has been switched off (no passive tracers).

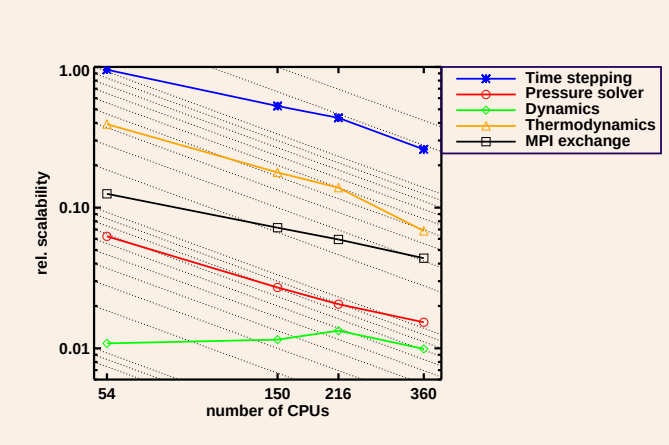


Figure 3. Same as Figure 1, but now including the biogeochemistry module and an optimized passive tracer communication.

behaviour of the model. Without the additional biogeochemical tracers the MPI exchange (and therefore the model as a whole) scales reasonably linear.

This showed that the MPI communication for the advection of the passive tracers had to be optimized. Instead of sending the data packages for the advection of each tracer over the halo boundaries of each model time individually, a new communication scheme was implemented that is adapted to the properties of the network, where the data for all passive tracers is exchanged jointly with one MPI_SEND/MPI_RECV statement. As Figure 3 shows, this decreased the time for the MPI exchange in the target model configuration with 216 processors significantly (by about 60%), which also resulted in an overall gain in efficiency.

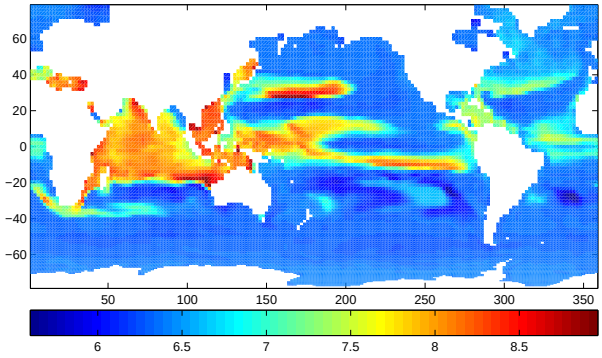


Figure 4. Modelled ratio between carbon and nitrogen (mol:mol) in the biomass within the oceanic surface mixed layer. At a sufficient supply of light and nutrients, phytoplankton cells tend toward a C.N ratio of around 6.6. The elevated values in the subtropical gyres are a consequence of strong nutrient limitation.

4 Results

Our main goal, a high-resolution run of the coupled physical-biogeochemical model over a decade, has not been reached during the first phase of the project due to the described scaling problems. We therefore just briefly demonstrate the potential of the model with some results from a coarse-resolution run and from the spinup phase of a high-resolution run.

The coarse-resolution model runs show that the representation of phytoplankton physiology in the model and the resulting decoupling of the elemental cycles of Si, C and N lead to new insights. In the subtropical gyres, which are biological 'deserts' due to the scarcity of nutrients, significantly elevated C:N ratios are modelled in the biomass (Figure 4). This would imply that these regions would contribute more that hitherto assumed to the vertical export of carbon into the deep ocean, given a fixed N and Si supply. This has consequences for the strength of the feedback between climate change and the marine carbon cycle through changes in stratification and therefore nutrient supply. The model also is able to reproduce the so-called 'opal-belt', a region of high deposition of diatom silica in the sediment around Antarctica, a feature that is usually not found in fixed-composition models (not shown).

At the same time, Figure 4 clearly shows deficits in the representation of the ocean circulation, such as an overestimate of nutrient upwelling along the equator. It will be very interesting to see whether the better reproduction in the high-resolution model of sharp fronts between nutrient-rich and nutrient-poor waters, e.g. in the ACC, will lead to an enhanced export of high C:N biomass into the deep ocean.

Model results from the high-resolution run (Figure 5) are not yet in equilibrium. However, they clearly show that the activity of eddies is reflected in the distribution of chlorophyll, especially along the ACC, and that the overly strong primary production around upwelling regions in the coarse-resolution run has been reduced and is much more in line

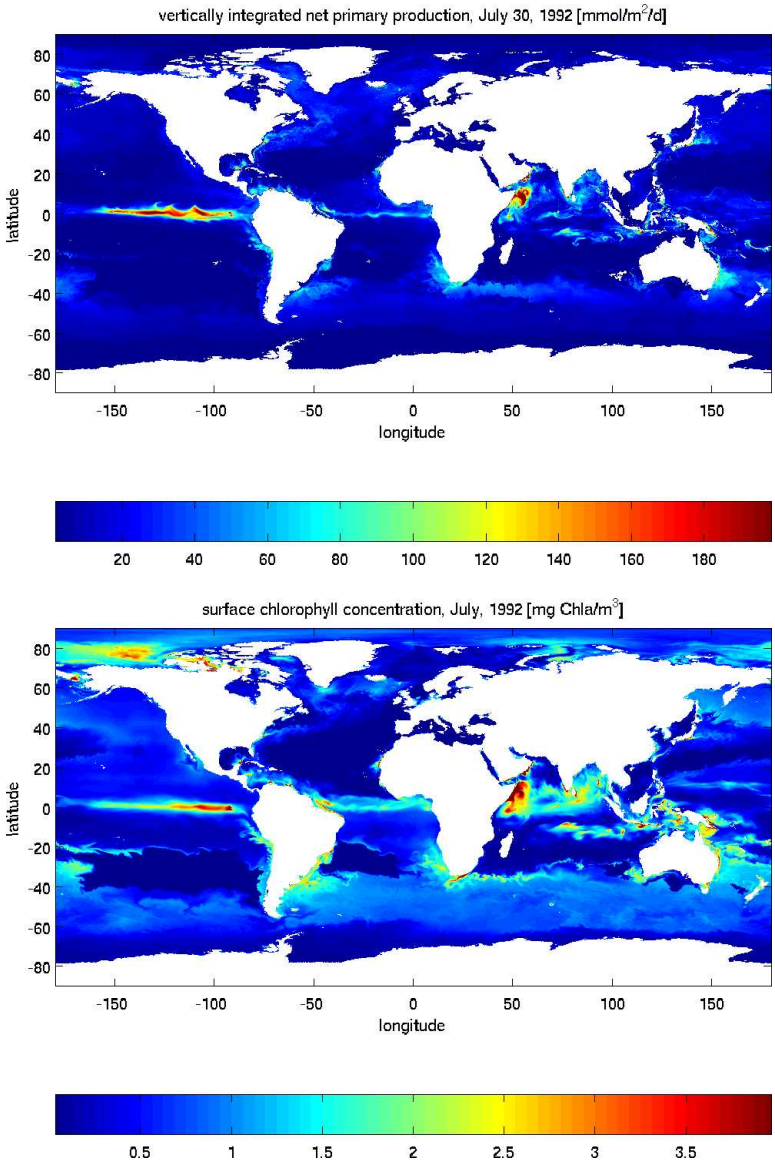


Figure 5. Results from a high-resolution run after 7 months of integration. Above: vertically integrated primary production on July 30, 1992. Below: mean concentration of chlorophyll-a at the sea surface for July, 1992.

with satellite-based estimates.

5 Concluding Remarks and Outlook

In spite of initial problems with scalability, the coupled MITgcm-REcoM model shown here shows some promise regarding the model-based description of the decoupling between the elemental cycles of carbon, nitrogen, silica and iron in a high-resolution model setup. Further model experiments will start from an optimized set of ecosystem parameter values that has been obtained by tuning the coarse-resolution model to observations of chlorophyll and export production, and will increasingly focus on the role of inter-annual variability on biogeochemical fluxes.

Acknowledgments

Markus Schartau (GKSS research centre, Geesthacht) has contributed much to the development of the biological model code. Sönke Hohn and Christoph Völker are supported by a EU grant within the EU FP6 program CarboOcean, contract number 511176 (GOCE). The benchmark computations were performed with a grant of computer time provided by the Research Centre Jülich.

References

1. E. Maier-Reimer, U. Mikolajewicz, and A. Winguth, *Future ocean uptake of CO_2: interaction between ocean circulation and biology*, Climate Dynamics, **12**, 711–721, 1996.
2. P.M. Cox, R. A. Betts, C.D. Jones, S. A. Spall, and I. J. Totterdell, *Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model*, Nature, **408**, 184–187, 2000.
3. P. Friedlingstein, J.-L. Dufresne, P.M. Cox, and P. Rayner, *How positive is the feedback between climate and the carbon cycle?*, Tellus, **55B**, 692–700, 2003.
4. J.C. Orr, E. Maier-Reimer, U. Mikolajewicz, P. Monfray, J.L. Sarmiento, J.R. Toggweiler, N.K. Taylor, J. Palmer, N. Gruber, C.L. Sabine, C. Le Queré, R.M. Key, and J. Boutin, *Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models*, Global Biogeochemical Cycles, **15**, 43–60, 2001.
5. J. C. Fyfe and O. A. Saenko, *Simulated changes in the extratropical Southern Hemisphere winds and currents*, Geophysical Research Letters, **33**, L06701, 2006.
6. G.M. Flato, *Sea-ice and its response to CO2 forcing as simulated by global climate models*, Climate Dynamics, **23**, no. 3-4, 229–241, 2004.
7. S.T. Gille, *Warming of the Southern Ocean since the 1950s*, Science, **295**, no. 5558, 1275–1277, 2002.
8. D. Olbers, D. Borowski, C. Völker, and J.-O. Wolff, *The dynamical balance, transport and circulation of the Antarctic Circumpolar Current*, Antarctic science, **16**, 439–470, 2004.
9. K.D. Six and E. Maier-Reimer, *Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model*, **10**, 559–583, 1996.
10. K. Fennel, M. Losch, J. Schröter, and M. Wenzel, *Testing a marine ecosystem model: Sensitivity analysis and parameter optimization*, Journal of Marine Systems, **28**, 45–63, 2001.

11. R.J. Geider, H.L. MacIntyre, and T.M. Kana, *A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature*, Limnology and Oceanography, **43**, no. 4, 679–694, 1998.
12. M. Schartau, A. Engel, J. Schröter, S. Thoms, C. Völker, and D.A. Wolf-Gladrow, *Modelling carbon overconsumption and the formation of extracellular particulate organic carbon*, Biogeosciences, **4**, 433–454, 2007.
13. J. Marshall, A. Adcroft, C. Hill, L. Perelman, and C. Heisey, *A Finite-Volume, Incompressible Navier Stokes Model for Studies of the Ocean on Parallel Computers*, Journal of Geophysical Research, **102**, no. C3, 5753–5766, 1997.
14. MITgcm Group, *MITgcm Release 1 Manual*, Online documentation, MIT/EAPS, Cambridge, MA 02139, USA, 2002, `http://mitgcm.org/sealion/online_documents/manual.html`.
15. W. G. Large, J. C. McWilliams, and S. C. Doney, *Oceanic Vertical Mixing: A Review and a Model with a Nonlocal Boundary Layer Parameterization*, Review of Geophysics, **32**, no. 4, 363–404, 1994.

Computer Science and Numerical Mathematics

Michael Griebel

Institut für Numerische Simulation, Universität Bonn
Wegelerstraße 6, 53115 Bonn, Germany
E-mail: griebel@ins.uni-bonn.de

Computer science aspects and numerical simulation on modern large scale parallel hardware platforms such as the JUBL are discussed in the following section. Here, we put the emphasis on the fast solution of large linear systems of equations by a parallel algebraic multigrid method, on the analysis of the scaling behaviour and the subsequent optimization of various software packages on the JUBL and on the performance-analysis tool SCALASCA and its properties to diagnose performance problems in large-scale parallel applications.

In their contribution *Coarse Grid Classification: AMG on Parallel Computers*, M. Griebel, B. Metsch, and M.-A. Schweitzer discuss their version of a parallel algebraic multigrid method (AMG) for the fast solution of large systems of linear equations. Here, in contrast to other algebraic multigrid methods, the creation of the coarse grids in the AMG approach is parallelized by means of a new coarse grid classification algorithm (CGC-ML). It is based on a graph agglomeration process using heavy edge matching, collapsing and merging edges in a multilevel fashion. The resulting parallel method is perfecly suited to very large numbers of processors and scales favourably on parallel computing systems like the JUBL.

In their contribution *Jülich Blue Gene/L Scaling Workshop 2006*, W. Frings, M.-A. Hermanns, B. Mohr, and B. Orth report on the main results of the first Blue Gene/L Scaling Workshop which took place in Jülich in December 2006. The areas of applications comprise the molecular dynamics study of radiation materials by the code DL_Poly3, the large scale *ab initio* DFT calculation of functional materials by the VASP program, the calculation of turbulent Rayleigh-Bénard convection by a pseudospectral code and numerical simulations of QCD by the BQCD package. Emphasis was put on the analysis of the scaling behaviour and the subsequent further optimization of these application packages on the JUBL.

In their contribution *Scalable Performance Analysis Methods for the Next Generation of Supercomputers*, F. Wolf, D. Becker, M. Greiner, and B. Wylie discuss the challenge which is posed by the actual trend in supercomputing architecture, i.e. using larger numbers of moderately fast processor cores rather than increasing the speed of uni-processors. Here, with an exponentially rising number of cores, the gap between peak performance and the performance actually sustained by production codes tends to widen even further. To overcome this problem, the performance-analysis tool SCALASCA has been designed. It is specifically suited for the use on large-scale systems such as the Blue Gene/L JUBL and its successor Blue Gene/P JUGENE and allows to identify wait states on systems with presently up to 16,383 processors. SCALASCA is explained in detail and it is demonstrated how this tool can be used to diagnose performance problems in large-scale parallel applications.

Coarse Grid Classification: AMG on Parallel Computers

Michael Griebel, Bram Metsch, and Marc Alexander Schweitzer

Institut für Numerische Simulation, Universität Bonn
Wegelerstraße 6, 53115 Bonn, Germany
E-mail: {griebel, metsch, schweitzer}@ins.uni-bonn.de

1 Introduction

Many phenomena in science and engineering can be described by a mathematical model. Such a model describes the relations between the variables and constants that are important for this problem, like velocities, forces or material properties. In many cases, these relationships are expressed by partial differential equations (PDEs).
Often it is not possible to obtain an analytical solution for these PDEs. Hence, one has to compute an approximate numerical solution. To this end, the underlying domain is discretized using a finite grid Ω. On this grid, the PDE is discretized using finite differences, finite elements of finite volumes. This results in a sparse (i.e. only few entries per row are different from zero) linear system of equations $Au = f$.
Multigrid Methods (MG)[1,6] are known to be optimal methods for such kind of linear systems. Algebraic Multigrid Methods (AMG)[2,10] extend the multigrid idea to a purely algebraic setting, i.e. the only input they need is the matrix A. They can be used if a geometric multigrid hierarchy is not available. This flexibility has a price: A *setup phase*, has to be carried out in which the multigrid hierarchy, i.e., the sequence of grids, transfer operators and coarse grid operators, is constructed automatically from A, before the well-known multigrid cycle (the *solution phase*) can be started.
For linear systems with millions of unknowns, parallel computation is necessary to speed up the computation and to be able to store the system in the memory. While the parallelization of geometric multigrid and hence the solution phase of algebraic multigrid is straightforward, the algorithms for creating the coarse grids in AMG are inherently sequential.
Various approaches for the parallelization of this step have been proposed over the years[7,8,3]. In this paper, we briefly describe the Coarse Grid Classification algorithm[4,5], as well as its extension CGC-ML to very large numbers of processors.

2 Algebraic Multigrid

In this section we give a short review of the algebraic multigrid (AMG) method. We consider a linear system $Au = f$, which comes from a discretization of a PDE on a grid $\Omega = \{1, \ldots, N\}$[a], with $A = (a_{ij})_{i,j=1}^{N}$ being a large sparse real matrix, $u = (u_i)_{i=1}^{N}$ and $f = (f_i)_{i=1}^{N}$ vectors of length N. We assume that A is a symmetric, positive definite M-matrix or an essentially positive matrix. To be able to solve this equation using the multigrid scheme[6], we need to specify the sequence of coarse grid operators A^l, transfer

[a] We denote grid points with their respective counting index in any dimension.

operators P^l and $R^l = (P^l)^T$, the smoothing operators S^l and the sequence of coarse grids Ω^l. To this end, in an algebraic multigrid method, we choose a simple smoothing scheme S^l, e.g. Gauss–Seidel or Jacobi relaxation. Then, we construct the grids $\{\Omega^l\}_{l=1}^L$, the transfer operators $\{P^l\}_{l=1}^{L-1}$ and the coarse grid operators $\{A^l\}_{l=1}^L$ in a recursive fashion depending on the fine grid operator $A = A^1$ only. This construction is carried out in the so-called setup phase which consists of three steps: the selection of an appropriate coarse grid

$$\Omega^{l+1} := C^l =: \Omega^l \setminus F^l,$$

the construction of a stable prolongation operator P^l, and the computation of the Galerkin coarse grid operator $A^{l+1} := (P^l)^T A^l P^l$. In the following, we omit the level index l where possible to simplify the notation.

Essential to many AMG schemes is the notion of a strong coupling between the grid point i and the grid point j. A grid point i is called *strongly coupled* to a grid point j if the corresponding matrix entry a_{ij} is relatively large. In the Ruge–Stüben coarsening (RSC) algorithm[9,10] for instance the coarsening (see Program 1) is based on the sets

$$S_i := \{j \neq i : \ -a_{ij} \geq \alpha \max_{k \neq i} |a_{ik}|\}, \ S_i^T := \{j \neq i : \ i \in S_j\}, \tag{1}$$

(typically $\alpha = 0.25$) which describe the strong connectivity graph of a matrix A. Along these strong couplings, the smooth parts of the error vary slowly. For an efficient coarse-grid correction of these parts, interpolation must also follow these couplings. This leads to three criteria for the choice of the coarse grid,

1. Let $i \in F$. Then for each $j \in S_i$ holds either $i \in C$ or $S_j \cap (C \cap S_i) \neq \emptyset$.
2. For all $i \in C$ and $j \in S_i \cup S_i^T$ holds $j \notin C$
3. C is a maximal set satisfying these properties.

In general, it is not possible to fulfill all these criteria. We choose to enforce the first one, as the stability of the interpolation depends on it. The coarsening process is carried out in

Program 1 AmgPhaseI (Ω, S, S^T, C, F)

```
begin
  U ← Ω;  C ← ∅;  F ← ∅;
  for i ∈ Ω do λ_i ← |S_i^T|; od;
  while max_{i∈U} λ_i ≠ 0 do
        i ← arg max_{j∈U} λ_j;                                   (*)
        C ← C ∪ {i};
        for j ∈ S_i^T ∩ U do F ← F ∪ {j};
            for k ∈ S_j ∩ U do λ_k ← λ_k + 1; od;
        od;
        for j ∈ S_i ∩ U do λ_j ← λ_j − 1; od;
  od;
  F ← F ∪ U;
end
```

two phases. In the first phase, see Program 1, an independent set C of coarse grid points is determined. Observe that we have a degree of freedom while choosing the first coarse grid point i (see the line indicated by $(*)$). After the first phase is carried out, a second

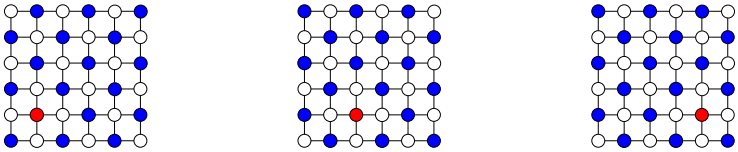


Figure 1. Resulting coarse grids for a 5-point discretization of the Laplace operator constructed by three different initial choices. The blue points indicate the respective coarse grid points, the red point indicates the first coarse grid point chosen.

coarsening phase checks that Criterion 1 is satisfied for all pairs $i, j \in F$ of fine grid points. If the condition $S_j \cap (C \cap S_i) \neq \emptyset$ is not fulfilled for a certain pair, then one of these points is added to the set of coarse grid points C.

Note that in Program 1 each point i chosen for the coarse grid C results in a change of the weights λ_j of all points j within two layers around the grid point i. This shows the sequential character of this coarsening algorithm, as the weight updates propagate throughout the whole domain during the coarsening loop. Hence, the parallelization of Ruge–Stüben-based AMG schemes is not straightforward.

3 Coarse Grid Classification

As can be seen from Figure 1, different choices for the first coarse grid point yield different coarse grids. There is no special advantage of using either one of these coarse grids in a sequential computation. In parallel computations, however, this gives us a degree of freedom to consistently *match* the coarse grids obtained on each processor individually by the RSC scheme at processor subdomain boundaries. This observation is the starting point for our coarse grid construction algorithm.

The CGC coarsening scheme employs a *two-stage* coarsening algorithm. First, we construct multiple coarse grids on each processor domain Ω_p independently by running the RSC algorithm multiple times with different initial coarse grid points. After the construction of these coarse grids on all processors, we need to select exactly one grid for each processor domain such that the union of these coarse grids forms a suitable coarse grid for the whole domain.

For details of the first stage, we refer to Program 2. Note that the number of iterations is bounded by the maximal number of strong couplings $|S_i|$ over all points $i \in \Omega_p$, which in turn is bounded by the maximal stencil width. Hence, the number of constructed grids ng_p per processor p is independent of the number of unknowns N and the number of processors P. The later constructed coarse grids may be of inferior quality but the selection mechanism described in the following will avoid them.
We now have obtained ng_p valid coarse grids $\{C_{(p),i}\}_{i=1}^{ng_p}$ on each processor p. To determine which grid to choose on each processor, we construct a directed, weighted graph $G = (V, E)$ whose vertices represent the candidate coarse grids,

$$V_p := \{C_{(p),i}\}_{i=1,\ldots,ng_p}, \quad V := \cup_{p=1}^{P} V_p.$$

The set of edges E consists of all pairs (v, u), $v \in V_p$, $u \in V_q$ such that $q \in \mathcal{S}_p$ is a

Program 2 CGC $(S, S^T, ng, \{C_i\}_{i=1}^{ng}, \{F_i\}_{i=1}^{ng})$

for $j \leftarrow 1$ **to** $|\Omega|$ **do** $\lambda_j \leftarrow |S_j^T|$; **od**;
$C_0 \leftarrow \emptyset$; $\lambda_{\max} \leftarrow \arg\max_{k\in\Omega} \lambda_k$;
do
  $U \leftarrow \Omega \setminus \bigcup_{i\le it} C_i$;
  if $\max_{k\in U} \lambda_k < \lambda_{\max}$ **then** $break$; **fi**;
  $it \leftarrow it + 1$; $F_{it} \leftarrow \emptyset$; $C_{it} \leftarrow \emptyset$;
  do
    $j \leftarrow \arg\max_{k\in U} \lambda_k$;
    if $\lambda_j = 0$ **then** $break$; **fi**;
    $C_{it} \leftarrow C_{it} \cup \{j\}$; $\lambda_j \leftarrow 0$;
    for $k \in S_j^T \cap U$ **do**
      $F_{it} \leftarrow F_{it} \cup \{k\}$; $\lambda_k \leftarrow 0$;
      for $l \in S_k \cap U$ **do** $\lambda_l \leftarrow \lambda_l + 1$; **od**;
    od;
    for $k \in S_j \cap U$ **do** $\lambda_k \leftarrow \lambda_k - 1$; **od**;
  od;
od
$ng \leftarrow it$;

(a) candidate coarse grids (b) consistent coarse grid

Figure 2. Discretization of the Laplace operator using finite differences. The left figure shows the candidate coarse grids indicated by different colours, the right figure shows the CGC graph.

neighbouring processor of p,

$$E_p := \{\cup_{q\in\mathcal{S}_p} \cup_{v\in V_p,\ u\in V_q} (v,u)\}, \quad E := \cup_{p=1}^{P} E_p,$$

where $\mathcal{S}_p$ is defined as the set of processors q with points j which strongly influence points i on processor p, i.e.

$$\mathcal{S}_p := \{q \neq p : \exists i \in \Omega_p,\ j \in \Omega_q :\ j \in S_i\}.$$

To each edge $e = (v,u)$ we assign a weight $\gamma(e)$ which measures the quality of the composed coarse grid if v and u are chosen to be part of it[4]. Figure 2 shows an example of the candidate coarse grids and the respective graph. After we have constructed the graph G of admissible local grids, we can use it to choose a particular coarse grid for each processor such that the union of these local grids automatically matches at subdomain boundaries. To this end, we transfer the whole graph onto a single processor. On this processor, we

Program 3 AmgCGCChoose $(V, H, \mathcal{C})$

```
begin
  C ← ∅; U ← V;
  for v ∈ U do λ_v ← |H_v| + |H_v^T|; od;
  for p ∈ {1, ..., P} do
      if λ_v = 0 for all v ∈ V_p
        then C ← {v | arbitrary v}; U ← U \ V_p; fi;
  od;
  while U ≠ ∅
        do
        v ← arg max_{w∈U} λ_w;
        C ← C ∪ {v}; U ← U \ V_p such that v ∈ V_p; λ_max ← max_{w∈U} λ_w;
        for w ∈ (H_v ∪ H_v^T) ∩ U do λ_w ← λ_max + 1; od;
  od;
end
```

choose exactly one node v_p from each subset $V_p \subset V$ with the following scheme: We first define heavy edges,

$$H_v := \cup_{q \in \mathcal{S}_p} \{w \,|\gamma(v,w) = \max_{u \in V_q} \gamma(v,u)\}, \text{ and } H_v^T := \{w \mid v \in H_w\}.$$

that indicate which candidate coarse grids can be fitted best to the coarse grid represented by $v \in V_p$. We then employ Program 3 to create the global coarse grid. This procedure takes up to P steps, one for each processor domain, see Program 3 for details. After running the algorithm, we transfer the choice $v_p \in \mathcal{C} \cap V_p$ back to processor p.

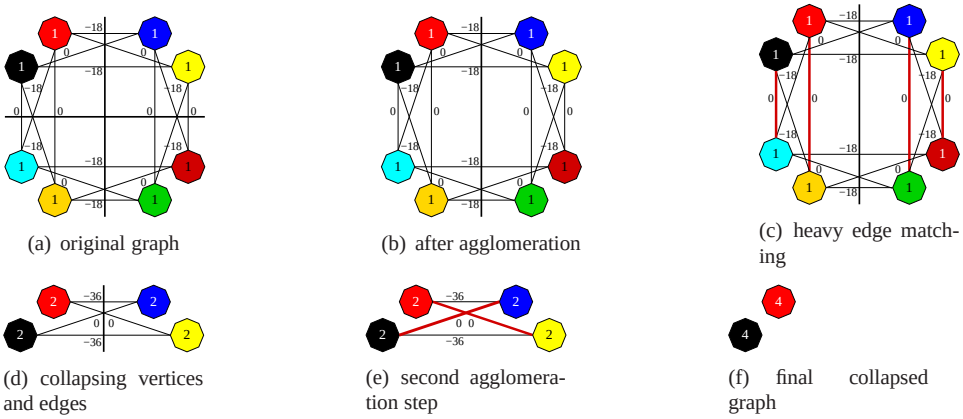
4 Multilevel Coarse Grid Classification

The main advantage of CGC over other parallel AMG is that the constructed coarse grids are very close to those a sequential AMG would produce. However, the original CGC has one major drawback: The graph representing the candidate coarse grids needs to be transferred to a single processor. For large numbers of processors ($P \gtrapprox 1000$) this leads to large communication costs as well as a significant run-time for the coarse grid selection algorithm AmgCGCChoose.

To overcome this issue, we have developed the CGC-ML algorithm which carries out the coarse grid selection procedure in a recursive, multilevel manner. We construct the graph $G = (V, E)$ as described in the last section. In addition, we assign a weight to each vertex which denotes the number of processor subdomains covered by the coarse grid represented by this vertex. Naturally, this weight is initialized with 1.

We do *not* transfer the whole graph onto a single processor. Instead, we proceed as follows:

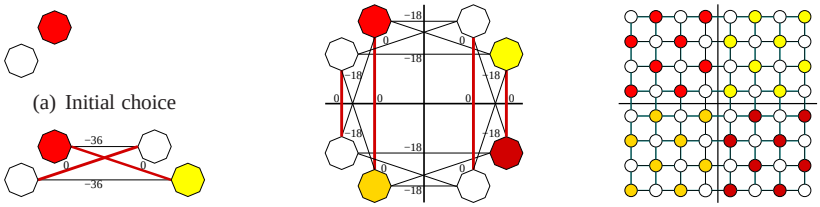
1. We agglomerate the graph on a subset of the processors, see Figure 3(b). Hence, a part of the edges (in this case, the vertical edges) do not cross the processor boundaries any more.
2. We can now employ a *heavy matching* on the inner edges of each processor subdomain, see Figure 3(c). Note that all edge weights are negative, so in fact the selected edge is the one with the smallest absolute weight.
3. We collapse the matched vertices and merge the edge sets, see Figure 3(d). Each vertex now represents a candidate coarse grid on an union of processor subdomains.



(a) original graph (b) after agglomeration (c) heavy edge matching

(d) collapsing vertices and edges (e) second agglomeration step (f) final collapsed graph

Figure 3. Graph clustering process. The graph is constructed by the CGC algorithm to a 5-point finite-difference discretization of the Laplace operator, distributed among four processors (cp. Figure 2(a)). The numbers in the vertices denote the number of subdomains covered by the coarse grid which is represented by the respective vertex. the number at each edge denotes the edge weight.



(a) Initial choice

(b) First refinement step (c) Second refinement step (d) Composed coarse grid

Figure 4. CGC-ML refinement process.

Accordingly, we update the vertex weights. The edge set of each vertex u is the union of the edge sets of the vertices v, w that were collapsed into u: $E_u \leftarrow E_v \cup E_w$. However, we never create an edge between two vertices which represent candidate coarse grids on the same processor subdomain. If two edges are collapsed into the same edge, we add their edge weights.

4. We proceed matching and collapsing the graph. If no further matching is possible, we again agglomerate the graph on a smaller subset of processors. If we are already on a single processor, we stop, see Figure 3(e) – 3(f).

We have now obtained a small set of vertices on a single processor. Now, we choose one vertex u such that it covers a maximal number of processor subdomains and mark it, see Figure 4(a). Then, we mark the vertices v and w that were collapsed into u. We recursively proceed refining this choice until we have reached the original graph, see 4(b) – 4(c). Now, on each processor subdomain, the candidate coarse grid represented by the marked vertex is selected as coarse grid for this processor subdomain and we obtain a coarse grid for the global discretization domain as depicted in Figure 4(d).

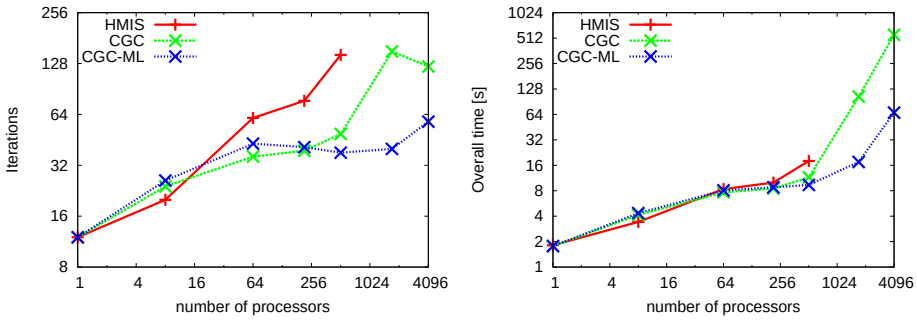


Figure 6. Number of iterations (left) and overall timings (right) for the model problem

During this refinement process, we must ensure that one vertex is marked per processor subdomain on each level of the graph hierarchy. In consequence, this will guarantee that after finishing the refinement, we have selected one candidate coarse grid on each processor subdomain. In our implementation, we proceed as follows. At each step in the refinement process where more processors are involved as in the previous step (i.e. a processor agglomeration was performed in the matching phase), we determine if a vertex is marked on *each* processor. If this is not the case, we mark the vertex that is most heavily coupled to the marked vertices on neighbouring processors.

Hence, on *each* level of the AMG multilevel hierarchy, we employ a multi-level graph coarsening algorithm.

5 Numerical Results

In this section, we present first numerical results obtained on the JUBL supercomputer. In particular, we employ AMG as a preconditioner for the conjugate gradient method and we compare the CGC-ML algorithm with the original CGC algorithm as well as the HMIS parallel coarsening algorithm[3].
We consider an model problem in three spatial dimensions,

$$-\nabla \cdot (a \nabla u) = f \tag{2}$$

on $[0,1]^3$ with Dirichlet boundary conditions. The diffusion coefficient a depends on (x, y, z) as depicted in Figure 5. We employ a 7-point finite difference scheme to discretize the problem on $31 \times 31 \times 31$ points per processor subdomain.

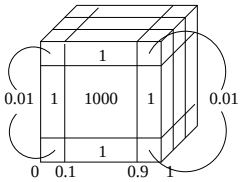


Figure 5. Values of the diffusion coefficient for Eq. 2

As strength threshold in (1), we set $\alpha = 0.25$. We omit the second coarsening pass of the RSC scheme as we use AMG as preconditioner only. Furthermore, we use the modified classical interpolation, see Ref. 7. On each level of the multigrid hierarchy, we employ a hybrid Gauss-Seidel/Jacobi smoother. We start the iterations with a zero initial vector u_0. The iteration is stopped if the residual $r_{it} = f - Au_{it}$ drops below 10^{-8} measured in the l^2-norm.

In Figure 6 we give the plots of the number of iterations and overall run-time for the considered parallel AMG schemes. From these plots we see that the CGC-ML algorithm achieves robust preconditioning for this problem up to thousands of processors. In contrast, the iteration numbers for HMIS coarsening increase significantly and the algorithm does not converge within 1000 steps for more than 512 processors. We see an increase of the total wall time requirements beyond 1024 processors which is caused by the slower communication between the racks. However, The CGC-ML algorithm shows a significantly improved scale-up behaviour compared with the original CGC algorithm.

Acknowledgments

This work was sponsored by the Sonderforschungsbereich 611, *Singuläre Phänomene und Skalierung in mathematischen Modellen*, sponsored by the *Deutsche Forschungsgemeinschaft*. We want to thank the Forschungszentrum Jülich for compute time on the JUMP and JUBL supercomputers.

References

1. A. Brandt, *Multi-level adaptive technique (MLAT) for fast numerical solution to boundary value problems*, In Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics, H. Cabannes and R. Teman, (eds.), pages 82–89, New York, Berlin, Heidelberg, Springer Verlag, 1973.
2. A. Brandt, S.F. McCormick, and J. Ruge, *Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations*, Institute for Computational Studies, POB 1852, Fort Collins, Colorado, 1982.
3. H. De Sterck, U. M. Yang, and J.J. Heys *Reducing complexity in parallel algebraic multigrid preconditioners*, SIAM J. on Matrix Analysis and Applications, **27**, pp. 1019–1039, 2006.
4. M. Griebel, B. Metsch, D. Oeltz, and M. A. Schweitzer, *Coarse grid classification: A parallel coarsening scheme for algebraic multigrid methods*, Numerical Linear Algebra with Applications, **13(2–3)**, pp.193–214, 2006.
5. M. Griebel, B. Metsch, and M. A. Schweitzer, *Coarse grid classification–Part II: Automatic coarse grid agglomeration for parallel AMG*, Preprint 271, Sonderforschungsbereich 611, Universität Bonn, 2006.
6. W. Hackbusch, *Multi-grid methods and applications*, Springer Series in Computational Mathematics. Springer-Verlag, Berlin, Heidelberg, 1985.
7. V. E. Henson and U. M Yang, *BoomerAMG: a parallel algebraic multigrid solver and preconditioner*, Applied Numerical Mathematics **41**, pp. 155-177, 2002.
8. A. Krechel and K. Stüben, *Parallel algebraic multigrid based on subdomain blocking*, Parallel Computing **27**, pp. 1009-1031, 2001.
9. J. W. Ruge and K. Stüben, *Algebraic multigrid (AMG)*, In Multigrid Methods, S. F. McCormick (Ed.), Frontiers in Applied Mathematics Vol. **5**, SIAM, 1986.
10. K. Stüben, *Algebraic multigrid (AMG): An introduction with applications*, In Multigrid, U. Trottenberg, C. W. Oosterlee, and A. Schüller, pages 413–532, Academic Press, 2001.

Jülich Blue Gene/L Scaling Workshop 2006

Wolfgang Frings, Marc-André Hermanns, Bernd Mohr, and Boris Orth

John von Neumann Institute for Computing
Jülich Supercomputing Centre
Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: {w.frings, m.a.hermanns, b.mohr, b.orth}@fz-juelich.de

From December 5-7, 2006, John von Neumann Institute for Computing, IBM, and the Blue Gene Consortium jointly held the first "Blue Gene/L Scaling Workshop" at Jülich. The aim of this workshop was to provide participants the chance to scale their codes across an 8-rack Blue Gene/L system. Besides the hardware, appropriate software and support personnel were provided to accomplish this task. Jülich provided about 800,000 CPU hours on the 16,384-processor Blue Gene/L system *JUBL* over a three-day period for the scaling runs. The participants of the workshop were selected by a peer review team. Selection criteria were (i) the confidence that the code would scale across 8 racks, (ii) the prospect that the JUBL infrastructure (OS, compilers, libraries) would meet the user's requests, and (iii) the scientific impact that the code could produce. Applicants selected were paired up with assigned advisors from Argonne National Laboratory, IBM, and NIC, who assisted in administrative (log-on, moving data, storing data) and scaling issues. This article highlights some of the results of running, scaling and optimising the participating applications on JUBL.

1 Introduction

In this article we summarise a selection of the results of the first "Blue Gene/L Scaling Workshop" which took place at Forschungszentrum Jülich from December 5-7, 2006. The aim of this workshop was to give both users and the support personnel from Argonne National Lab, IBM, and JSC/NIC the opportunity to investigate the performance and scaling behaviour of selected applications on JUBL, the 8-rack IBM Blue Gene/L system in Jülich.

Unfortunately, the limited space here does not allow for a complete account of all the valuable results and experiences gained during the three days of the workshop. Besides those participants whose results are described below, others have contributed equally to the workshop's big success. Among them were A. Dubey from the University of Chicago, who investigated the performance and scaling of the FLASH code from the field of astrophysics, and A. Nishida from Chuo University, Japan, who studied the scalability of a numerical software library for large scale scientific simulations. The performance analysis of the CFD package XNS with SCALASCA, a tool developed at JSC, proved especially successful for M. Nicolai and M. Probst from the group of M. Behr at RWTH Aachen University. Not only did it result in a significant improvement of the code, it also marked the beginning of a still ongoing, fruitful collaboration between this group and the parallel performance analysis group at JSC.

The following sections describe the outcome of running, scaling and optimising the four remaining participating applications. They are based on the participants' own workshop reports, collected in a JSC Technical Report[1]. For further details, also on the applications omitted here, please consult this reference and the references therein.

2 Molecular Dynamics Studies of Radiation Hard Materials

DL_POLY3 is a classical molecular dynamics package developed at CCLRC Daresbury Laboratory, from where it was brought to the workshop by I. J. Bush. Prior to the workshop, the code had never been run on systems with significantly more than 1000 processors. The system investigated at the workshop was a model of radiation damage in a fluoritized Zirconium pyrochlore. One of the native Gadolinium ions in the system was replaced by Uranium, which was then given a velocity consistent with a 100 keV recoil after an alpha decay. Due to the very high velocity of the Uranium ion it was necessary to study very large supercells, and the total system size used was approximately 14.6 million particles. The power of the Jülich Blue Gene/L allowed, for the first time, to model this system with a realistic recoil.

The DL_POLY3 code is a totally distributed memory code which uses a link-cell algorithm that scales approximately like $O(N)$. The various terms of the force field can be generalised as (i) short range repulsion, (ii) Van Der Waal's (VDW) attraction, and (iii) Coulomb forces.

The first two are both short range terms, and are handled together in DL_POLY3. Subsequent references to VDW terms therefore include both these terms. Due to the short range of these forces they are expected to scale very well with processor count due to their spatial locality (compare halo exchange algorithms). On the other hand, Coulomb forces are long range terms, and have to be handled differently. The Ewald sum technique used in DL_POLY3 splits the evaluation into two terms – one short ranged, one long ranged. While the former is evaluated in real space (very similar to the VDW terms), the long range term has to be handled differently. DL_POLY3 uses the Smooth Particle Mesh Ewald (SPME) algorithm, the key feature of which is a Fast Fourier Transform (FFT). For this DaFT is used, a package written at Daresbury Laboratory that is novel in that it avoids performing 'all-to-all' operations by parallelising the individual 1D FFTs.

Results

Once ported, the code ran and scaled very well "out of the box". The good scaling for MD of the test system is shown in Fig. 1. All jobs were run in virtual node mode.

Of the various elements of the force field the VDW and short range Ewald terms both scale almost perfectly, and at least for VDW terms the expected deviation from ideal behaviour at 16384 processors is not very apparent. As expected, the long range Ewald terms, i.e. those terms that require an FFT, scale less well. However, given the comparatively small size of the FFT grid, $512 \times 512 \times 512$, the scaling is still good.

The absolute times per time step for each of the components of the execution showed that, at these processor counts and for this system, the dominant term was the long range component of the Ewald summation. The most important result, however, was that the time for an MD time step on 16,384 processors was sufficiently small to allow full simulations to be performed in a realistic amount of time.

The one major problem that was experienced was I/O, which was prohibitively slow. This was due to the fact that, at the time of the workshop, all I/O in DL_POLY3 was performed in serial, i.e. all through one processor, and to/from formatted files. The reasons for this were simplicity and portability of the files, and that the time taken for I/O had

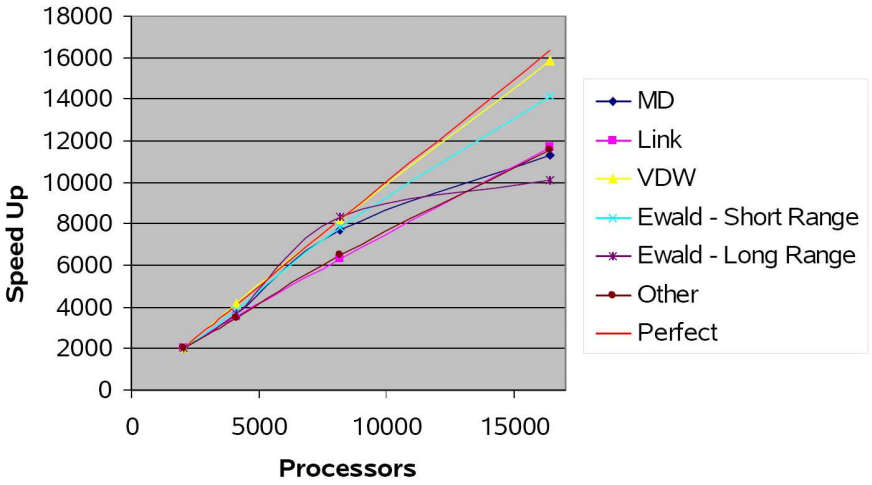


Figure 1. Scaling of DL_POLY3 on JUBL. ”MD” shows the scaling of the total computation time, ”Link” refers to the time taken at each time step to implement the link-cell algorithm (and build the Verlet neighbour list), and the remainder are the components of the force field terms outlined above.

always been small compared to the compute time. However, it became very clear during the workshop that the large system sizes enabled by the massively parallel Blue Gene/L require new, parallel ways of implementing I/O.

Summary

DL_POLY3 was demonstrated to scale well out to 16,384 processors on Blue Gene/L. Porting was straightforward. It was shown that the code runs fast enough on 8 racks to allow a detailed scientific study of the system. Two candidates for optimisation which were found using Xprofiler and explicit timings were the loop nest that interpolates the ionic charges onto the regular grid, and the message passing time in the FFT.

3 Large Scale *Ab Initio* Calculations of Functional Materials

Ab initio Density Functional Theory (DFT) codes have become a widely used tool for the investigation of unusual materials properties and the prediction of novel functional materials. Unlike empirical and semi-empirical methods their accuracy is not hampered by simplifying assumptions about the interatomic interactions, so that their use becomes inevitable where electronic and structural properties become closely interrelated, as it is often the case for modern functional materials. Off-the-shelf DFT codes like VASP (Vienna *Ab initio* Simulation Package), which was investigated at the workshop by M. E. Gruner from the group of P. Entel at the University of Duisburg-Essen, were historically not designed to run large problems on a vast number of processors. However, a complete redevelopment appears infeasible due to their complexity. The readiness of existing codes for modern massively parallel supercomputer architectures is therefore pivotal for the question whether *ab initio* materials science on the mesoscopic scale will become possible in near future.

In previous tests it could be shown that the code scales well up to 1,024 nodes. However, further scaling was hampered by parts of the code which make heavy use of the SCALAPACK eigensolver `pdsyevx` and the parallel 3d Fast Fourier Transforms (FFT). One strategy to circumvent this limitation is the tackling of larger systems, where the relative amount of time spent for interprocess communication will be reduced. The option to optimise the communication structure in the code was not projected at this stage, but may become a target for future efforts. Therefore, three questions were at the centre of interest during the scaling workshop: (i) Is the VASP code capable of running on 8k nodes? (ii) How severe is the limitation of 512 MB per node for larger problems on several thousand nodes? (iii) What is the maximum system size that can be tackled on the Blue Gene/L?

To answer these questions, realistic test systems were simulated. The first system consisted of super-cells of the magnetic shape alloy Ni_2MnGa of various sizes (576, 672, 720, 768, 800, 896 and 1024 atoms), sufficiently large to contain a martensitic twin boundary. These twin boundaries can be shifted in realistic magnetic fields which gives rise to the so-called ferromagnetic shape memory effect, which makes these alloys interesting for a new class of magneto-mechanical actuators. In addition, calculations of large super-cells of the dilute magnetic semiconductor ZnO:Co and GaN:Gd were performed. Three different system sizes were categorised, consisting of a total number of 432, 864 and 1296 atoms. In both cases, due to the large super-cells, integration in k-space was restricted to the Γ-point.

Results

During the workshop it became evident that only 512 MB of memory per node represent a severe limitation. To alleviate this restriction, the coprocessor mode was used throughout all of the calculations.

The largest system that could be simulated on one and two racks was a Ni_2MnGa supercell consisting of 672 atoms (6720 spin-polarised valence electrons). Here, a speedup of 1.73 between one and two racks could be achieved. The largest system that could be tested successfully on the full 8 racks was an 800 atom Ni_2MnGa supercell comprising 8000 spin-polarised valence electrons. Here, a few self-consistency steps could be achieved, which gave enough data for a timing analysis. On a cube (2×2 racks) this problem was brought to complete self-consistency. The measured speedup between four and eight racks, however, was only 1.22. Reasonable timing data could not be obtained for two racks and less due to memory restrictions. However, for a previous test case (561 iron atoms), which was not hampered by memory limitations down to 128 nodes, a speedup of 1.31 between one and two racks was achieved, while on 1024 nodes an efficiency of 70% of the ideal performance, extrapolated from 128 nodes, was measured. It appears plausible to expect an overall efficiency of above 50% for the 800 atom super-cell on the cube. For the doped ZnO the largest system calculated consisted of 864 (7725 valence electrons) atoms which could be computed on one rack.

Summary

From the experience gained at the workshop it was concluded that on a Blue Gene/L system with 512 MB per node the maximum partition that can be used efficiently is four racks.

The largest problems that were manageable with VASP on such an installation contained 800-900 atoms and up to 8000 (spin-polarised) valence electrons. On Blue Gene (L or P) systems with 1 GB per node larger systems can be treated, and efficient scaling will probably be achieved even on 8 or 16 racks. Improvements in the FFT communication scheme, which are planned in close cooperation with IBM, might result in significant performance gains.

4 Turbulent Convection with Very Large Aspect Ratios

Turbulence often occurs in geometries where the lateral dimensions exceed the vertical dimension by orders of magnitude. For example, atmospheric mesoscale layers typically have lateral extensions of up to 1000 km and are characterised by aspect ratios $L : W : H = 1000 : 1000 : 1$.

At the workshop, J. Schumacher from the Technical University of Ilmenau studied turbulent Rayleigh-Bénard convection, which represents one particular process in such mesoscale layers. His three-dimensional pseudospectral simulation code advances the Boussinesq equations for an incompressible fluid in time by means of a second-order predictor-corrector scheme. Lateral boundary conditions are periodic; vertical boundary conditions are free-slip.

One of the core parts of the numerical scheme used in this simulation is the Fast Fourier Transform (FFT). The classical parallel implementation of three-dimensional FFTs uses a slab-wise decomposition of the simulation domain. For a simulation with N^3 grid points, the method allows a parallelisation on up to N processors. Due to the rather small memory size per core, the Blue Gene/L requires a *volumetric* FFT which decomposes the three-dimensional volume into cuboid-like rods and hence allows a parallelisation degree of N^2.

The prime requirement for a simulation with a large grid is that the subdomains of the grid (including buffers and temporary storage) fit into the 512 MB of memory on a single Blue Gene/L node. At the same time the FFT algorithm should of course be scalable, i.e. increasing the number of CPUs to solve the problem should also substantially decrease the time-to-answer.

Results

Three FFT packages were compared during the workshop: the old slab-wise method, the BGL3DFFT by M. Eleftheriou *et al.*, and the P3DFFT package by D. Pekurovsky. After some tests the P3DFFT package turned out to be the best solution in terms of performance. Moreover, its interface and implementation met the application's needs in an optimal way (real-to-complex/complex-to-real).

In addition to the inclusion of the P3DFFT package, further optimisations contributed to the improvement of the code performance. These include the optimisation of the process topology, loop-splittings to enable code vectorization in the calculation of the r.h.s. of the Boussinesq equations, and the use of non-blocking point-to-point communication instead of MPI_Reduce and MPI_Bcast.

Further analysis indicated that with these improvements almost all of the remaining communication overhead is related to calls to MPI_Alltoallv required by the FFT algorithm. They are thus the only limiting factor for strong scaling. Future improvements

CPUs	processor mode	$iproc \times jproc$	time
1024	co	32×32	205.7 s
2048	vn	32×64	183.3 s
4096	co	64×64	90.66 s
4096	vn	128×32	109.6 s
4096	vn	64×64	118.6 s
8192	co	128×64	60.5 s
8192	co	512×16	74.0 s
8192	vn	128×64	77.8 s
16384	vn	128×128	62.6 s

Table 1. Runtime tests for the aspect ratio $\Gamma = 32$. The CPU time was measured with MPI_Wtime() for a loop of 5 time steps.

of the implemenation of MPI_Alltoallv will therefore have an immediate impact on the performance of this application.

Table 4 summarizes the findings for a grid resolution intended for production runs: $N_x \times N_y \times N_z = 4096 \times 4096 \times 256$ at $\Gamma = 32$. For such a large grid it turned out to be necessary to implement some internal summation loops in double precision.

The parameters $iproc$ and $jproc$ in the table represent the two dimensions of the processor grid that has to be generated for the volumetric FFTs. The differences in execution time for coprocessor or virtual node mode are small for large problem sizes, which is due to the fact the local subdomains no longer fit into the L3 cache of a Blue Gene/L node, so that the two cores compete for the memory bandwidth.

Summary

The use of the P3DFFT package and a number of further improvements in the communication overhead improved significantly the performance of this code on Blue Gene/L. These efforts allowed to run production jobs on at least two racks. Since the code performance is bound by the memory bandwidth for problem sizes of interest, VN mode does not result in a significant speedup compared to CO mode.

5 Numerical Simulations of QCD

The objective of current projects of the QCDSF collaboration is to better constrain the extrapolation of lattice results to the chiral and continuum limits by performing simulations at more realistic quark masses and at smaller lattice spacings. This has become possible due to substantial improvements in the Hybrid Monte Carlo (HMC) algorithm and a significant increase in computing power.

The program investigated by H. Stüben from ZIB Berlin is called BQCD (Berlin QCD). It is a Hybrid Monte Carlo code that simulates quantum chromodynamics (QCD) with Wilson gauge action and non-perturbatively $O(a)$ improved Wilson fermions. The program is written in Fortran 90 and uses MPI. For Blue Gene/L the most compute intensive part has been implemented in assembler.

Lattice QCD is defined on a four-dimensional regular lattice with (anti-)periodic boundary conditions. The kernel of BQCD is a standard conjugate gradient solver with even/odd preconditioning. Typically, 80 % of the execution time is spent in this solver. The dominant operation is the matrix-vector multiplication. In the context of QCD the matrix is called *fermion matrix*. This fermion matrix is sparse with eight entries per row. The entries in row i are the nearest neighbours of entry i of the vector.

The entries of the fermion matrix are 3×3 complex matrices and the entries of the vector are 3×4 complex matrices. Experience on current machines shows that the performance of the Fortran code for the matrix-vector multiplication is about 10 % of peak, and approximately 20 % in case of the assembler version. On the Hitachi SR8000-F1, which was one of the main production machine for the QCDSF collaboration from 2000–2006, the Fortran code ran at about 40 % of peak.

Parallelism of the program is achieved using domain decomposition. In the parallelised *cg*-kernel ghost cells of the input vector have to be exchanged before the vector can be multiplied with the fermion matrix. In order to scale QCD programmes up to high numbers of processes an excellent communication network is required because the lattice volume per core becomes small. On the other hand small lattice volumes per core improve the utilisation of data caches and hence improve performance.

For the floating point operations of the fermion matrix multiplication the SIMD *(double hummer)* instructions of the floating point units are used, which can perform four floating point operations per clock cycle. The communication is done in parallel with the computation, so that in principle computation and communication can overlap. For the communication between nearest neighbours in three of the four directions the torus network is used. The network is accessed from the nodes via memory-mapped fifos. Because of the simple communication pattern (of nearest neighbour data exchange only), no dynamical routing is necessary. Furthermore, each node receives the data packets in the order in which they are needed, so that no reordering of packets is necessary. Each node sends a packet to its neighbour one iteration before the neighbouring node needs the data, so that the network latency can be hidden.

In the fourth direction, which corresponds to the x-direction of the lattice, the lattice is split between the two CPUs in one node, so that the communication can be done via the shared memory. However, since the L1 caches are not coherent, the straightforward approach of simply reading data from the other CPU's memory does not work. At present, a memory region in the L3 cache (the scratchpad area) is used, which is marked L1-caching inhibited for data exchange between the two cores.

Results

In scaling tests during the workshop the performance of the *cg*-kernel was measured. To this end the code was instrumented with timer calls, and for the kernel all floating-point operations were counted manually.

In order to get good performance it is important that the lattice fits the physical torus of the machine. In order to achieve this, MPI process ranks have to be assigned properly. On Blue Gene/L this can be accomplished by setting the environment variable `BGLMPI_MAPPING` appropriately.

The results of the performance tests are illustrated by double logarithmic plots in Fig. 2 (the dotted lines indicate linear scaling). As can be seen from the plots the Fortran/MPI

version exposes super-linear scaling on the $48^3 \times 96$ lattice. Even the $32^3 \times 64$ lattice scales quite well given the fact that the lattice volumes per core are tiny. For the same tiny local lattices the scaling of the assembler version is even better than that of the Fortran/MPI version. This means that in the assembler version computation and communication really overlap.

5.1 Summary

At the workshop it was found that the BQCD code scales up to the full 8 racks of the Jülich Blue Gene/L. The highest performance measured was 8.05 TFlop/s on the whole machine. This high performance could be obtained by using *double hummer* instructions and techniques to overlap communication and computation.

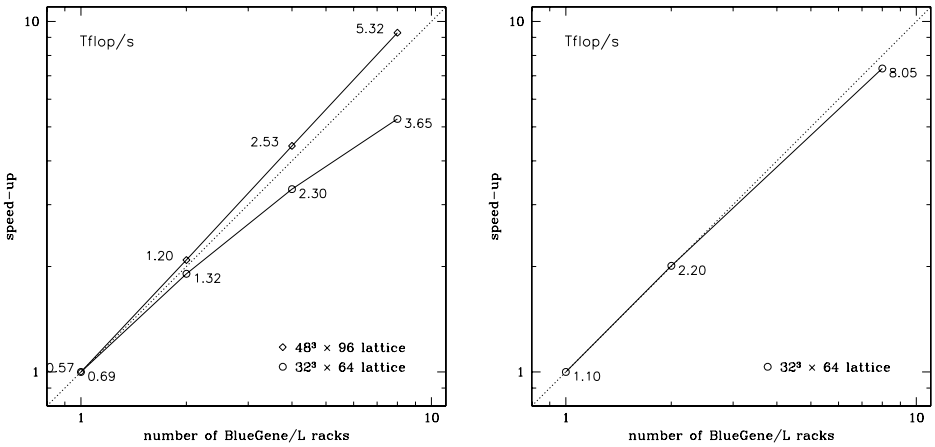


Figure 2. Scaling of the *cg*-kernel of BQCD in the Fortran 90/MPI version (left), and in the assembler version (right).

Acknowledgments

The authors would like to thank their co-organizers from IBM and the Blue Gene Consortium, all participants, and last but not least all their colleagues from JSC who helped making this workshop a success.

References

1. W. Frings, M.-A. Hermanns, B. Mohr, B. Orth (Eds.) *Blue Gene/L Scaling Workshop 2006*, Technical Report IB-2007-02, 2007. For an online version see `www.fz-juelich.de/jsc/files/docs/ib/ib-07/ib-2007-02.pdf`.

Scalable Performance Analysis Methods for the Next Generation of Supercomputers

Felix Wolf[1,2], **Daniel Becker**[1,2], **Markus Geimer**[1], **and Brian J. N. Wylie**[1]

[1] Forschungszentrum Jülich, Jülich Supercomputing Centre
52425 Jülich, Germany
E-mail: {f.wolf, d.becker, m.geimer, b.wylie}@fz-juelich.de

[2] RWTH Aachen University, Department of Computer Science
52056 Aachen, Germany

Facing increasing power dissipation and little instruction-level parallelism left to exploit, computer architects are realizing further performance gains by using larger numbers of moderately fast processor cores rather than by further increasing the speed of uni-processors. As a consequence, supercomputing applications are required to harness much higher degrees of parallelism in order to satisfy their growing demand for computing power. However, writing code that runs efficiently on large numbers of processors remains a significant challenge.

To address this challenge, the Helmholtz-University Young Investigators Group *Performance Analysis of Parallel Programs* at the Jülich Supercomputing Centre develops performance-analysis tools to diagnose inefficiencies in supercomputer applications and works with application developers to analyze and improve the performance of their codes. In this contribution, we highlight the research activities of our group during the past two years and give an outlook on future work. At the centre of our report lies the development of SCALASCA, a performance-analysis tool that has been specifically designed for large-scale systems and that allows the automatic identification of harmful wait states in applications running on thousands of processors.

1 Introduction

Supercomputing is a key technology pillar of modern science and engineering, indispensable to solve critical problems of high complexity. The extension of the ESFRI road map to include a European supercomputer infrastructure in combination with the creation of the PRACE consortium acknowledges that the requirements of many critical applications can only be met by the most advanced custom-built large-scale computer systems. However, as a prerequisite for their productive use, the HPC community needs powerful and robust development tools. These would not only help improve the scalability characteristics of scientific codes and thus expand their potential, but also allow domain scientists to concentrate on the underlying models rather than to spend a major fraction of their time tuning their application on a particular machine.

As the current trend in microprocessor development continues, this need will become even stronger in the future. Facing increasing power dissipation and little instruction-level parallelism left to exploit, computer architects are realizing further performance gains by using larger numbers of moderately fast processor cores rather than by further increasing the speed of uni-processors. As a consequence, supercomputer applications are being required to harness much higher degrees of parallelism in order to satisfy their growing demand for computing power. With an exponentially rising number of cores, the often substantial gap between peak performance and the performance level actually sustained

by production codes is expected to widen even further. Finally, increased concurrency levels place higher scalability demands not only on applications but also on parallel programming tools. When applied to larger numbers of cores, familiar tools often cease to work in a satisfactory manner (e.g., due to escalating memory requirements or limited I/O bandwidth).

To overcome this challenge, the Helmholtz-University Young Investigators Group *Performance Analysis of Parallel Programs* at the Jülich Supercomputing Centre develops performance-analysis tools to diagnose inefficiencies in supercomputer applications running on thousands of processors and works directly with application developers to improve the performance of their codes. The main product of the group is the software package SCALASCA[1], a comprehensive performance-analysis tool that has been specifically designed for use on large-scale systems, such as the Blue Gene/L JUBL in Jülich and its successor Blue Gene/P JUGENE. When scaling message-passing applications to thousands of processors, their performance is often affected by wait states that occur when processes fail to cooperate efficiently. Building upon earlier experience obtained during the KOJAK project[2], SCALASCA is able to efficiently identify such wait states at the previously inaccessible scale of 16,384 processes and shows potential even for larger configurations.

In this article, we highlight the research accomplishments of our group during the past two years. We explain SCALASCA in detail and demonstrate how it can be used to diagnose performance problems in large-scale parallel applications. We also make a side trip to computational grids and demonstrate that SCALASCA's novel architecture proves beneficial there as well. At the end, we outline our research goals for the coming years.

2 SCALASCA

In message-passing (i.e., MPI) applications, which still constitute the major portion of large-scale applications running on systems such as IBM Blue Gene or Cray XT, processes often require access to data provided by remote processes, making the progress of a receiving process dependent upon the progress of a sending process. As a consequence, a significant fraction of the time spent in communication and synchronization routines can often be attributed to wait states that occur when processes fail to reach implicit or explicit synchronization points in a timely manner, for example, as a result of unevenly distributed workloads. Especially when trying to scale communication-intensive applications to large processor counts, such wait states can present severe challenges to achieving good performance. As a first step in reducing the impact of wait states, application developers need a diagnostic method that allows their localization, classification, and quantification especially at larger scales. Because wait states cause temporal displacements between program events occurring on different processes, their identification can be accomplished by logging those events with timestamps into so-called *event traces* and then searching these traces for patterns indicating situations where a process waits for input from one or more other processes.

2.1 Scalability

Compared to SCALASCA's predecessor KOJAK, the novel approach taken in SCALASCA is that the event traces are searched in a much more scalable way by exploiting both distributed memory and parallel processing capabilities available on the target system. Instead

of sequentially analyzing a single global trace file, as KOJAK does, SCALASCA analyzes separate process-local trace files in parallel by *replaying* the original communication on as many CPUs as have been used to execute the target application itself. Since trace processing capabilities (i.e., processors and memory) grow proportionally with the number of application processes, we can achieve good scalability at previously intractable scales. In brief, to meet the scalability requirements of next-generation supercomputers, SCALASCA is a parallel program in its own right.

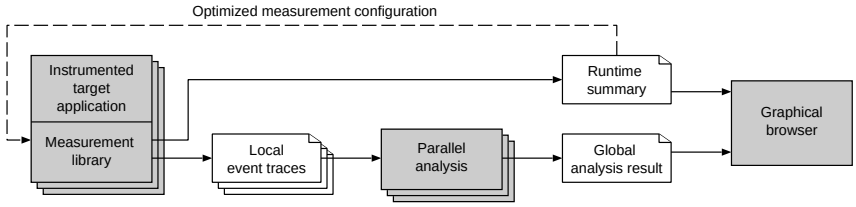


Figure 1. Parallel trace analysis in SCALASCA.

Figure 1 shows the basic analysis workflow carried out by SCALASCA. Before any performance data can be collected, the target application must be *instrumented*, that is, extra code must be inserted to record the events whenever they occur. On some systems including Blue Gene, this can be done completely automatically using compiler support; on other systems a mix of manual and automatic instrumentation mechanisms is offered. When running the instrumented code on the parallel machine, the user can choose between two options: (i) generating a runtime summary with aggregate performance metrics for individual function-call paths or (ii) generating event traces to record individual runtime events. The first option is useful to get an overview of the performance behaviour and also to optimize the instrumentation for later trace generation. Since traces tend to become very large, this step is usually recommended before choosing the second option. When tracing is enabled, each process generates a trace file containing records for all its process-local events. After program termination, SCALASCA loads the trace files into main memory and analyzes them in parallel using as many CPUs as have been used for the target application itself. During the analysis, SCALASCA classifies detected pattern instances by category and quantifies their significance for every function-call path and system resource involved. Both results of the trace analysis as well as runtime summaries can be interactively explored in a graphical browser (Figure 5).

To demonstrate the scalability of the parallel analysis approach taken in SCALASCA, we compared it to the sequential approach taken in KOJAK. The ASCI SMG2000 benchmark, a parallel semi-coarsening multigrid solver that uses a complex communication pattern, served as a test case. SMG2000 performs a large number of non-nearest-neighbour point-to-point communication operations and can be considered to be a stress-test for the network subsystems of a machine. Applying a weak scaling strategy, a fixed $64 \times 64 \times 32$ problem size per process with five solver iterations was configured, resulting in a nearly constant application runtime as further CPUs were added.

Figure 2 charts wall-clock execution times for the uninstrumented benchmark and the analyses of trace files generated by an instrumented version with a range of process num-

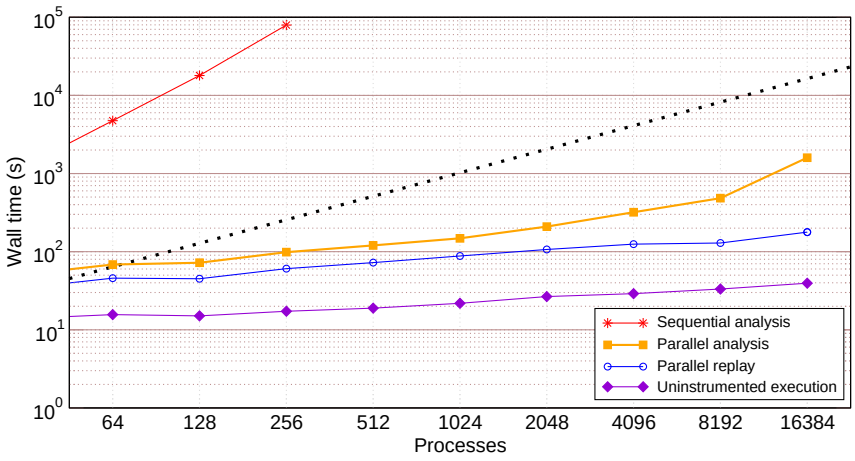


Figure 2. Comparison of application execution, sequential trace analysis (KOJAK) and parallel trace analysis (SCALASCA) times for the ASCI SMG2000 benchmark on up to 16,384 processors of Blue Gene/L. Linear scaling is represented by the bold dotted line.

bers on JUBL. The 8-fold doubling of process numbers necessitates a log–log scale to show the corresponding range of times, particularly for the old sequential analysis (which furthermore becomes impractical for the largest traces). The figure shows the total time needed for the parallel analysis and the time taken by the parallel replay itself without file I/O. It can be seen that the total analysis time including loading the traces into the analyzer never exceeded 30 minutes.

The actual procedure of replaying and analyzing the event traces without file I/O exhibits smooth scaling behaviour up to very large configurations. Because of its replay-based nature, the time needed for this part of the analysis depends on the communication behaviour of the target application. Since communication is a key factor in the scaling behaviour of the target application as well, similarities can be seen in the way both curves evolve as the number of processes increases. Notably, the total time for the new analysis approach is orders of magnitude faster than the sequential analysis based on KOJAK even at modest process counts, making it possible to examine traces at previously intractable scales in a reasonable time.

2.2 Clock Synchronization

Identifying wait states in event traces of message-passing applications requires measuring temporal displacements between concurrent events, although many parallel systems, such as PC clusters, do not provide synchronized hardware clocks. In these cases, linear interpolation techniques can already account for differences in offset and drift, assuming that the drift of an individual processor is not time dependent. However, inaccuracies and drifts varying in time can still cause violations of the logical event ordering that are harmful to the accuracy of our analysis. The *controlled logical clock* algorithm by Rabenseifner[3] compensates for such violations in point-to-point communication by shifting message events in time as much as needed while trying to preserve the length of intervals between lo-

cal events. Our group extended this method to collective communication to enable a more complete correction of realistic message-passing traces. In addition, we designed a parallel version of the algorithm that scales to thousands of application processes.[4]

2.3 Computational Grids

If a single machine does not provide enough CPUs to solve a given problem, multiple independent parallel machines can be combined into a so-called *metacomputer* that appears to the application as a single coherent system. However, achieving satisfactory application performance on such a metacomputer is hard, because high latency of inter-machine communication as well as differences in hardware of constituent machines may introduce various types of wait states. Since the analyses offered by SCALASCA could prove especially beneficial in such a grid-like environment, we extended[5] our tool in such a way that it can cope with typical metacomputer limitations, such as a missing global file system and varying network latencies. In addition, we added metacomputing-specific patterns to SCALASCA's pattern base. Using this grid-enabled version of SCALASCA, we were able to remove harmful wait states from an environmental-science application running on the VIOLA[6] grid testbed.[7]

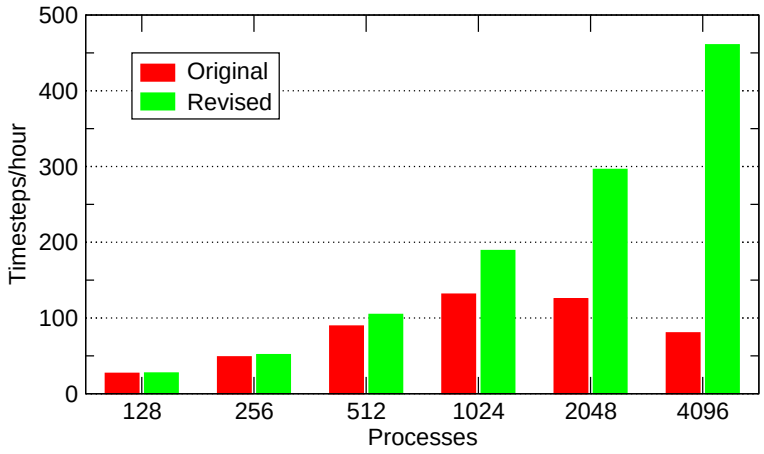


Figure 3. Performance improvement of XNS after removing redundant message traffic.

3 Application Engagement

In addition to engineering performance tools, our group also actively works with application developers to analyze and improve the performance of their codes. A recent example is the XNS fluid dynamics application being developed at the Chair for Computational Analysis of Technical Systems at RWTH Aachen University. XNS can be used for effective simulations of unsteady fluid flows, including microstructured liquids, in situations involving significant deformations of the computational domain, as they occur in blood pumps

deployed to support human heart function. The algorithm is based on finite-element techniques on irregular three-dimensional meshes[8].

Initially, XNS was successfully running on JUBL, however, its scalability was unsatisfactory and resisted scaling beyond approximately 900 processes until the first Jülich Blue Gene Scaling Workshop in December 2006 provided an opportunity for the application developers and our group to start working together. After a first investigation of the solver using basic profiling tools already hinted toward redundant message traffic (i.e., zero-sized messages), a statistical trace analysis using the SCALASCA infrastructure showed that the number of zero-sized messages were rapidly growing with the number of processes employed. Subsequent remediation[9] allowed the application to continue scaling with a four-fold simulation performance improvement at 4,096 processes (Figure 3), demonstrating the benefits of interdisciplinary collaboration between domain scientists and performance analysts.

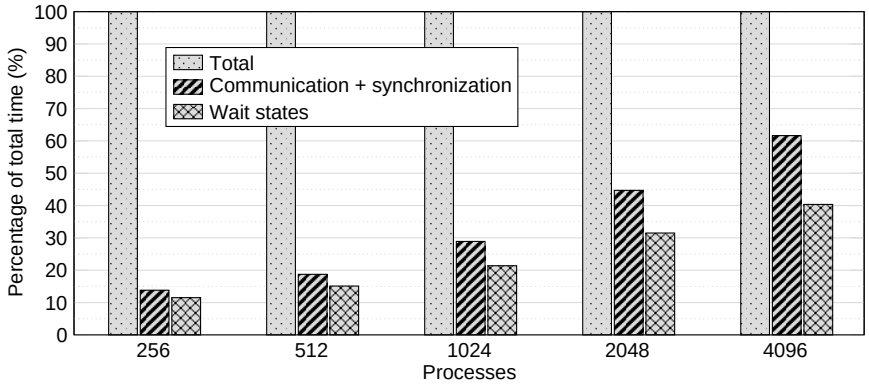


Figure 4. Performance behaviour of XNS (tuned version) at a range of scales from 256 to 4096 processes. The bars show the percentage of time the application spent in communication & synchronization operations including associated wait states.

Continuing our investigation, we applied our trace analysis to the tuned version, which exhibits almost perfect scaling behaviour on up to 512 processors. However, as the number of processors is raised further, the parallel efficiency continuously degrades, although even at the largest configuration of 4096 processes a clear speedup can still be observed (Figure 3). Beyond 512 processes, as we can see in Figure 4, the communication and synchronization overhead grows steeply. Yet the primary result of our analysis is that the biggest fraction of this overhead is actually waiting time, in the 4096 CPU case amounting to roughly 40 % of the total time (at least 25 % when accounting for intrusion overhead), illustrating that the wait states we are targeting can constitute principal performance problems at larger scales. The SCALASCA display in Figure 5 shows a function-call path that appears to be a major source of the wait states diagnosed during our analysis. Of course, with this finding we hope that the performance of XNS can be further improved.

4 Outlook

Potential causal connections between different wait state instances or related phenomena, such as load imbalance, are currently not covered in a systematic way. Understanding such connections, however, could prove essential for more effective scaling strategies. One approach of establishing links between different wait-state instances would be to define hypotheses and subsequently verify them using a trace-based simulator. Similar in spirit to approaches such as Dimemas[10], the simulator could leverage SCALASCA's parallel archi-

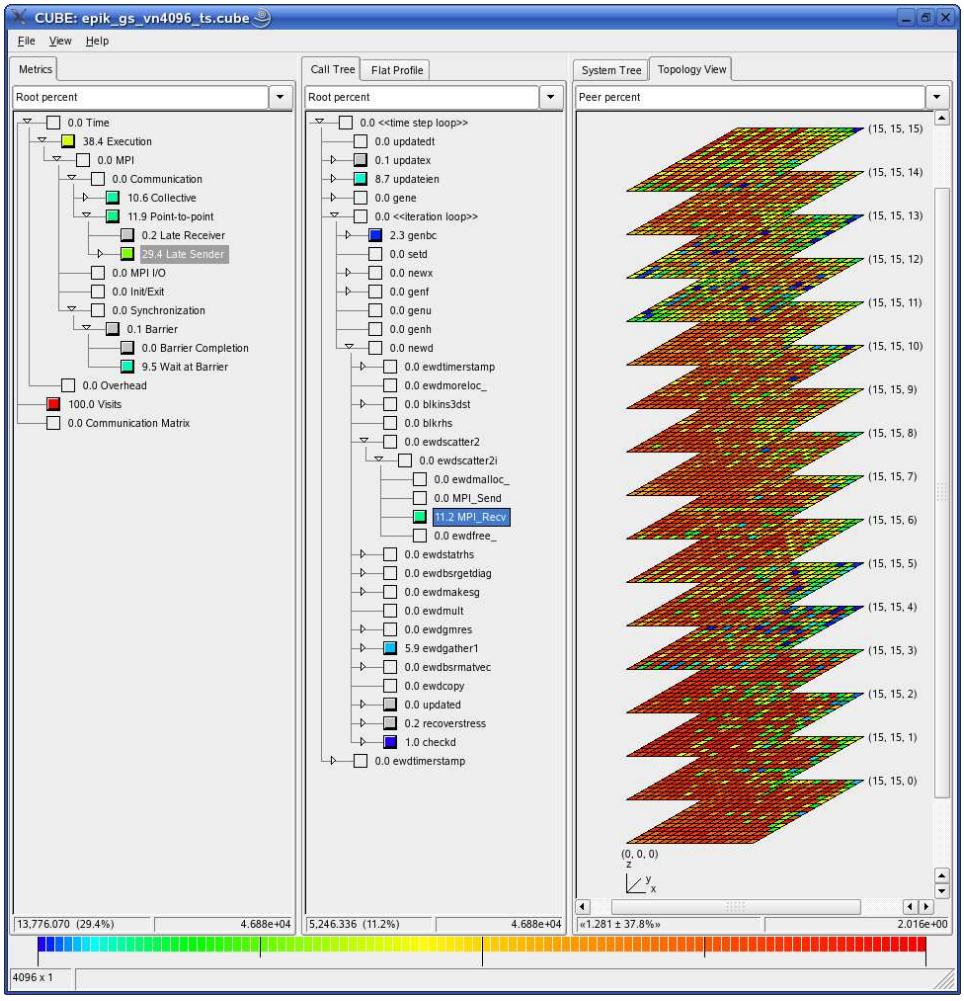


Figure 5. XNS (tuned version) wait states with 4096 processes in `ewdscatter2()`. The middle pane shows the distribution of waiting times across the call tree as the percentage of the time spent in the time step loop. The right pane visualizes how the time incurred by the selected call path is spread across the physical Blue Gene/L torus topology.

tecture to achieve the required scalability.

To institutionalize our major national and international collaborations on HPC programming tools, we recently founded the Helmholtz Virtual Institute - High Productivity Supercomputing together with the Technical University of Dresden, RWTH Aachen University, and the University of Tennessee. The mission of VI-HPS is devoted to the development and deployment of advanced and integrated performance analysis and error detection tools for high-performance computing applications. A significant portion of our resources is also assigned to support and training activities. We expect that this initiative will further promote the quality and accelerate the development process of complex simulation codes in science and engineering.

References

1. Markus Geimer, Felix Wolf, Brian J. N. Wylie, and Bernd Mohr, *Scalable Parallel Trace-Based Performance Analysis*, in: Proc. 13th European PVM/MPI Users' Group Meeting, vol. 4192 of *LNCS*, pp. 303–312, Springer, Bonn, Germany, September 2006.
2. F. Wolf and B. Mohr, *Automatic performance analysis of hybrid MPI/OpenMP applications*, Journal of Systems Architecture, **49**, no. 10-11, 421–439, 2003.
3. R. Rabenseifner, *Die geregelte logische Uhr, eine globale Uhr für die tracebasierte Überwachung paralleler Anwendungen*, PhD thesis, Universität Stuttgart, March 2000.
4. D. Becker, R. Rabenseifner, and F. Wolf, *Timestamp Synchronization for Event Traces of Large-Scale Message-Passing Applications*, in: Proc. of the 14th European Parallel Virtual Machine and Message Passing Interface Conference (EuroPVM/MPI), vol. 4757 of *LNCS*, pp. 315–325, Springer, Paris, France, September 2007.
5. D. Becker, F. Wolf, W. Frings, M. Geimer, B. Wylie, and B. Mohr, *Automatic Trace-Based Performance Analysis of Metacomputing Applications*, in: Proc. of the International Parallel & Distributed Processing Symposium (IPDPS), IEEE Computer Society, Long Beach, CA, March 2007.
6. BMBF (Ministry for Education and Research), *Vertically Integrated Optical Testbed for Large Applications in DFN (VIOLA)*, `http://www.viola-testbed.de/`.
7. D. Becker, W. Frings, and F. Wolf, *Performance Evaluation and Optimization of Parallel GRID Computing Applications*, in: Proc. of the 16th Euromicro Workshop on Parallel and Distributed Processing (PDP), Toulouse, France, February 2008, (to appear).
8. M. Behr, D. Arora, O. Coronado, and M. Pasquali, *Models and Finite Element Techniques for Blood Flow Simulation*, International Journal of Computational Fluid Dynamics, **20**, 175–181, 2006.
9. B. J. N. Wylie, M. Geimer, M. Nicolai, and M. Probst, *Performance analysis and tuning of the XNS CFD solver on BlueGene/L*, in: Proc. of the 14th European Parallel Virtual Machine and Message Passing Interface Conference (EuroPVM/MPI), vol. 4757 of *LNCS*, pp. 107–116, Springer, Paris, France, September 2007.
10. J. Labarta, S. Girona, V. Pillet, T. Cortes, and L. Gregoris, *DiP : A Parallel Program Development Environment*, in: Proc. of the 2nd International Euro-Par Conference, vol. 1124 of *LNCS*, pp. 665–674, Springer, Lyon, France, August 1996.

Hydrodynamics and Turbulence

Peter Grassberger

Department of Physics & Astronomy, University of Calgary
2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4
E-mail: p.grassberger@fz-juelich.de

Although the contributions to this section all deal with turbulence and with the instabilities of spatially extended systems, they cover a wide range of subjects. They reflect well the importance of turbulence, both in technological applications and in pure science. Turbulence has been since the 19th century the prototype of an inherently nonlinear phenomenon. While some understanding had been achieved in seminal works by Helmholtz, Kelvin, Prantl, Kolmogorov, Heisenberg, and others, it had been understood from very early on that the best one can reach is a statistical description. And even that is far from easy.

While turbulence (and hydrodynamics in general) had always been an important subject in engineering and in applied sciences, it had been less important in the physics curriculum after the revolution brought about by quantum mechanics. With the recent emphasis on nonlinear phenomena and with the ability to study them numerically on electronic computers, the emphasis on turbulence is also back in physics. Not the least this is also due to the power of modern computational tools, without which the modern revival of turbulence would not have been possible.

Although there are only four contributions to this section, they illustrate well the range of problems and of motivations encountered in turbulence studies. Let us first discuss the paper on ventricular assist devices by M. Behbahani *et al.* It deals with blood flow in artifical organs, i.e. with hydrodynamic turbulence in medicine. This reflects the increased importance of biomedical technology. It might not seem as big a challenge as the flow around an aircraft or inside a combustion engine, but it poses similar difficulties as far as complexity of the geometry and non-stationarity of the flow boundaries are concerned. An added problem is the frailty of red blood cells which requires a delicate balance between efficiency and smoothness of the flow. It also requires a detailed understanding of the flow down to the smallest length scales. As the authors show, their simulations can be done efficiently on the Jülich Blue Gene/L supercomputer, but they also show that the full power of modern state-of-the-art supercomputing is needed.

The next paper, by H. Homann *et al.*, is on plasma turbulence, i.e. on magnetohydrodynamics. This is of relevance for nuclear fusion reactors and for astrophysics, and it poses even more challenges due to the added degrees of freedom. Although the authors restrict themselves to rather simple boundary conditions, they are still haunted by geometrical complexities, this time due to the need for adaptive mesh refinements. One particular achievement of their simulations is that they are able, for the first time, to see a clear scaling regime in the Eulerian framework. As in the previous case, this would not have been possible without the massively parallel supercomputers in Jülich.

Understanding the scaling structure of fully developed hydrodynamic turbulence is the subject of the third paper, by J. Schumacher *et al.*. This deals with the central theoretical

question, of whether there are universal scaling laws similar to those which are at the core of the theories of critical phenomena and of relativistic quantum field theory. Due to the absence of a usable perturbation theory, analytic progress has been much slower in turbulence than in these other fields, and all recent achievements rely on the careful analysis of wet lab and numerical "experiments". Although the Reynolds (resp. Rayleigh) numbers reached in the simulations cannot yet compete with those reached in real experiments, this is compensated by the accessibility of noise-free and precise measurements on even the finest scales. Accordingly, simulations like those discussed in the paper by Schumacher *et al.*, which use the most sophisticated modern hardware, represent a priceless window into this terra incognita.

The last of the four papers, "Lyapunov Instabilities of Extended Systems" by H.-L. Yang *et al.*, is somewhat different from the others. It does not deal with turbulence proper, but with the slowest response of spatially extended chaotic systems to perturbations. In a chaotic system the response to a perturbation typically grows exponentially in time, with different Lyapunov exponents for different perturbations. If such a system is translationally invariant, one expects the 'Lyapunov modes' with slowest growth to be related to hydrodynamic modes. This is indeed seen in simulations and in analytic calculations, but details had previously been obscured by the difficulty to simulate large enough systems for long enough time. It seems that this is now overcome by the simulations of Yang *et al.*. In particular, they do not confirm a strange difference between hard and soft core fluids claimed by previous authors. Unfortunately, however, this chapter is not yet closed. As shown recently by F. Ginelli *et al.* (arXiv:0706.0510), the entire concept of Lyapunov modes has to be reconsidered. It might be that the results obtained so far survive unchanged, but it might also be that we are in for major surprises.

Towards Shape Optimization for Ventricular Assist Devices Using Parallel Stabilized FEM

Mehdi Behbahani, Marek Behr, Mike Nicolai, and Markus Probst

Chair for Computational Analysis of Technical Systems (CATS)
Center for Computational Engineering Science (CCES)
RWTH Aachen University
Steinbachstr. 53B, 52074 Aachen, Germany
E-mail: {mb, behr, nicolai, probst}@cats.rwth-aachen.de
http://www.cats.rwth-aachen.de

Over the last decade, computer simulations of fluid flow have steadily gained acceptance as an effective tool for evaluation of design modifications. The flow features in complex geometries such as blood pumps, as well as the evolution of those features resulting from design changes, are hard to predict even by experienced design engineers. Linking a suitable mathematical framework with appropriate models in order to evaluate shape modifications will ultimately allow not only to analyze flow features but also to compute an optimal design directly. Such a framework requires an immense amount of computing power, which makes it a perfect candidate to exploit the potential of parallel processing on high-performance systems such as the Jülich Blue Gene/L.

1 Introduction

Diseases of the heart are a leading cause of death in the industrialized nations. The most reliable therapy for end-stage heart failure—heart replacement via a transplant—can be applied only in a fraction of the cases because of dramatic shortage of suitable donor hearts. Since 1960's, attempts are being made to design a mechanical solution to heart failure; such a solution can take the form of a full replica of the heart—dual pumping chambers and complex valves—or, more commonly, of an assisting device, which pumps the blood from the existing failing ventricle into the aorta. The latter is referred to as Ventricular Assist Device, or VAD.

The Chair for Computational Analysis of Technical Systems (CATS) at the RWTH Aachen University, under the direction of Prof. M. Behr, is specializing in Computational Fluid Dynamics (CFD) analysis and has been working on simulation of blood flow in VADs since 2000, with the latest analyses focusing on the miniature MicroMed DeBakey VAD, see Fig. 1. The DeBakey pump is an axial pump that provides a countinuous flow. Its design renders it suitable for both long-term and pediatric use.

2 Simulation

The computational flow analyses described here are performed with XNS, an in-house CFD code for simulations of unsteady fluid flows in situations involving significant deformations of the computational domain. The underlying equations are the incompressible Navier-Stokes equations. Although solving these equations analytically remains out of reach, there has been considerable progress in the development of numerical methods that can provide

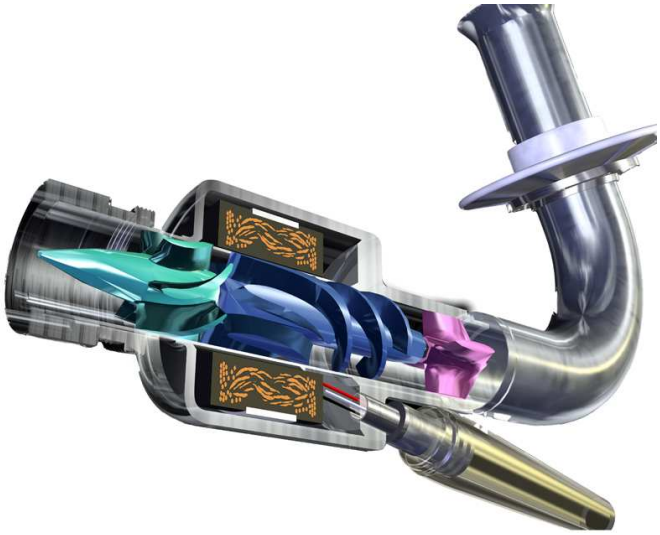


Figure 1. DeBakey VAD geometry. From right to left: inflow cannula, straightener, impeller and diffuser.

reliable approximations of the solution. XNS uses stabilized finite element (FE) formulations together with a space-time discretization on unstructured three-dimensional meshes. Important features for simulating the flow in a VAD include the ability to handle rapidly translating or rotating boundaries using the shear-slip mesh update method (SSMUM), as well as simulating flows of microstructured (in particular viscoelastic) liquids.

Simply computing the flow field inside the VAD presents a formidable challenge regarding the required computing power. Resolving the flow characteristics appropriately necessitates a high-resolution mesh with millions of elements, see Fig. 4. The use of hundreds or even thousands of processors is inevitable; the pump's impeller operates at speeds up to 10,000 rpm necessitating simulations that span thousands of time steps.

In the following sections, we report on two crucial aspects in the design of VADs that can be investigated by numerical simulations. As the most basic requirement, the pump has to be able to operate against a certain pressure head to provide a desired flow rate of 2–5 l/min. Another concern is biocompatibility after implantation; the flow patterns should not induce blood cell damage (hemolysis) above a critical level where it can cause multiple-organ failure, and recirculation zones should be minimized to prevent blood clotting (thrombosis).

Hydraulic Performance

The function of a blood pump is to provide the human body with the required blood flow volume of typically 5 l/min for an adult. Simulations aim to predict the performance of a pump under operating conditions. In order to confirm numerically obtained results it is possible to compare them to measured values from in-vitro experiments, i.e., experiments in a closed mock-up flow loop using an original blood pump.

From a point of view of pump performance it is important to determine the pressure

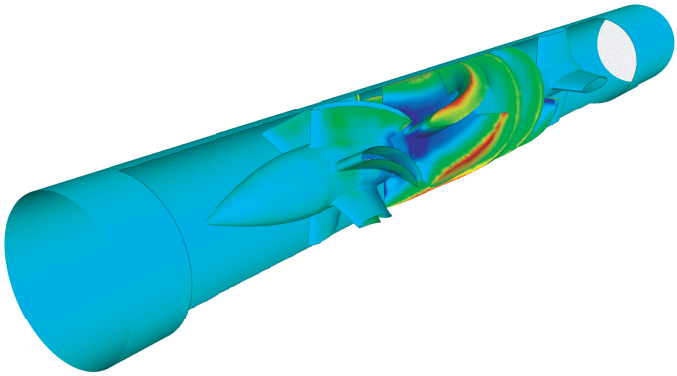


Figure 2. Sample pressure distribution on DeBakey VAD.

head (pressure difference between in- and outflow) and the flow rate at a given rotational speed of the impeller. In the case of the Debakey pump the Particle Image Velocimetry (PIV) method was used to achieve data points which define the hydraulic performance curves for rotational speeds of 7500, 8500 and 9500 rpm. First data points obtained from numerical simulations show good agreement with the experimental curves at a rotational speed of 7500 rpm where computed pressure heads at a given flow volume are within 5% of experimental values. A visualization of the pressure distribution on the surface of the pump can be seen in Fig. 2. The red colour on the top of the impeller vanes indicates high pressure due to the small gap between vanes and housing. At 9500 rpm the agreement is not so good with computed head being about 80% of the experimental one; one possible reason for this could be imperfections in the turbulence model used.

A new simulation series on the Blue Gene with an improved geometry model is now being conducted. We typically use 2048 processors, which permits to compute one pump revolution in about 3 hours. For visualization purposes, it may be necessary to store all of the computed pressure and velocity values for one revolution, corresponding to data amounts of up to 30 GB. To receive a quasi-stationary flow condition up to twenty revolutions can be required. By mesh refinement it can be confirmed if a reduction of the mesh size has no further influence on the outcome of the computations, which is a prerequisite for convergence. Former simulations were performed with a 5.5 million element mesh and the new series uses 7.5 million elements.

Blood Damage

Blood is a highly-complex fluid, a suspension of red and white blood cells (RBC, WBC) and platelets in a plasma. The volume fraction of RBC is about 45%, and blood damage, or hemolysis, can be caused by prolonged elevated stresses acting on those cells.

With increasing levels of shear stress the coin-stacked arrangement of the RBCs breaks up, the membrane starts rotating around the encapsulated fluid until, eventually, its hemoglobin contents are partly released into the plasma through small pores. This is illustrated in the upper sequence of Fig. 3. Up to a certain amount this so-called plasma-free hemoglobin can be filtered by the kidneys; higher concentrations lead to intoxication and,

in a worst-case scenario, death.

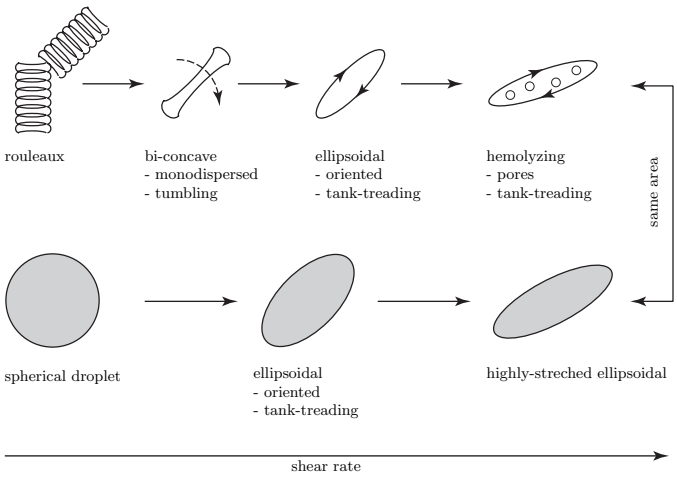


Figure 3. Behaviour under shearing: RBCs (top), droplet model (bottom).

Our group uses a novel approach to assess hemolysis. Instead of simply using exposure time and the flow-induced shear stress, the RBCs are modelled as droplets. As shown in the lower sequence in Fig. 3, hemolysis is initiated when the increase of the surface area of the droplet (which takes an ellipsoidal shape under shearing) and of a hemolyzing RBC match. Tracking the deformation along flow pathlines enables us to capture important physiological properties of the cells, such as membrane rotation (tank-treading) and relaxation in low shear regions.

Since these phenomena contribute to reduction of hemolysis, our deformation-based model showed good correspondence with experimentally-diagnosed hemolysis values[1] while the conventional stress-based models are known to overpredict hemolysis.

3 Parallel Computing

As already mentioned in Sec. 2, efficient simulations of the DeBakey VAD rely strongly on high-performance computing on powerful machines like the 8-rack Blue Gene/L system at Research Centre Jülich. The parallel implementation of XNS is based on message-passing communication libraries and exploits graph-based mesh-partitioning techniques. This means that the computational mesh is divided into subsets of elements which are then assigned to the available processors. A sample partitioning of the DeBakey blood pump is shown in Fig. 4.

After distributing the elements among the processors, a discrete form of the Navier-Stokes flow equations is solved simultaneously on all processors. While parts of the computation can be carried out within a single partition, others require communication between the processors and the exchange of data buffers. Since this inter-processor communication

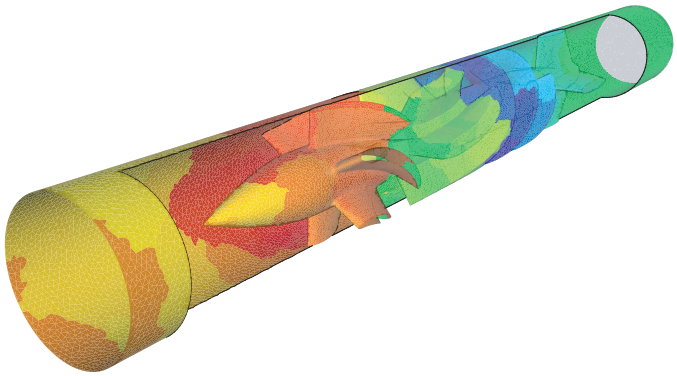
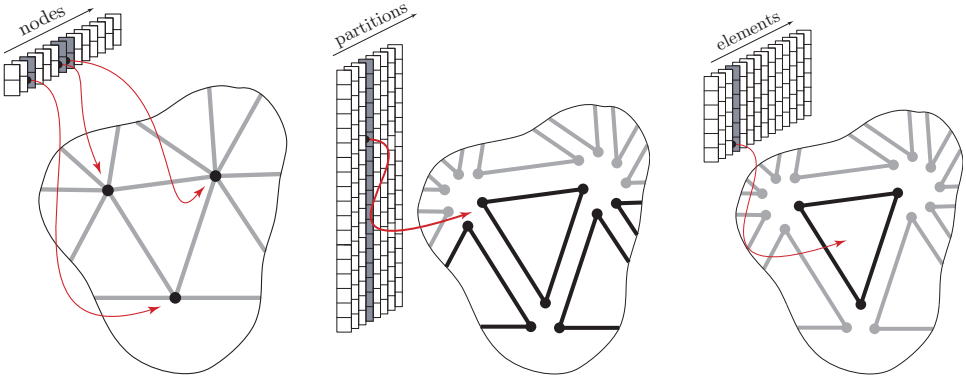


Figure 4. DeBakey blood pump: element partitions for 1024 processors.

is limited to the subdomain boundaries, the element subsets should be assigned in a way that reduces the overall boundary size to a minumum.

There are also different ways of storing the flow information during the simulation, each of which can be advantageous depending on the operation that has to be performed with the data. Fig. 5 gives an impression about the storage modes used in XNS for a triangular two-dimensional mesh. The nodal data is assigned to the processors according to the element partitioning and can be used to obtain element-based values. Moving from one level to another requires information about the element distribution as well as a connectivity matrix that specifies for each element the nodes it is composed of.



(a) node-based (b) partition-based (c) element-based

Figure 5. Storage modes for finite element data.

Performance on Jülich Blue Gene/L

In December 2006, a scaling workshop[2] for applications running on the 8-rack Blue Gene/L system in Jülich was organized and sponsored jointly by the John von Neumann Institute for Computing (NIC), IBM and the Blue Gene Consortium. Prior to the workshop, one could observe an acceptable scaling of XNS up to 1024 processors while there was no significant speed-up above that (see Fig. 6). To find the bottlenecks in the code, the communication between the processes during the simulation runs was analyzed, both with XNS internal time measurements and the SCALASCA package[3]. After improving the communication patterns of XNS, the simulation performance could be improved remarkably up to 4096 processes (see Fig. 6). A good scaling is expected also for 8192 processes; this is to be analyzed in future test runs.

The possibility to make efficient use of up to one fourth of the processors available on the Jülich Blue Gene/L is of great value for the simulations, because it allows to generate the required data in a reasonable time. Going towards the ultimate goal of the DeBakey project – the computation of an optimal design with respect to criteria mentioned in Sec. 2 – one could profit from the good scalabilty of XNS and make efficient use of even bigger machines.

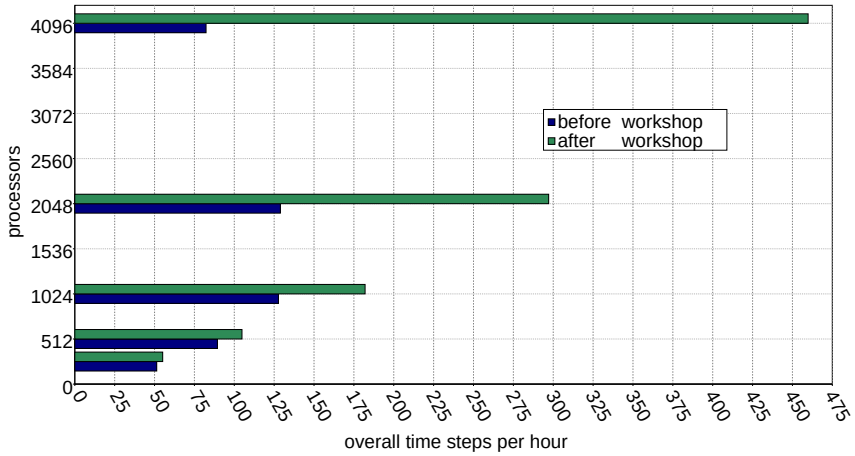


Figure 6. Performance of XNS on Jülich Blue Gene/L.

As outlined in Sec. 4, the computational demands and memory requirements in design optimization are much higher compared to plain simulations since the Navier-Stokes equations have to be solved several times on different geometries.

4 Shape Optimization for Fluids

As already mentioned, we aim to use an optimization framework that is capable of achieving an optimal design in an iterative, but automatic way, i.e., without a need for user interaction. Using such a framework, the user only needs to specify areas of the geometry, for example the impeller of the pump, that should be subjected to optimization and define

a so-called objective function to judge the quality of a design modification. This can be a difficult task: for blood pumps, an objective function that relates to hemolysis seems most reasonable, but it is important that a gradient of this function can be computed which again limits one's choice.

The state-of-the-art approach is to couple a flow solver with a grid generation tool which are both driven by an optimization algorithm. This approach has several drawbacks. Normally it is not possible to get the exact gradient in such an environment; therefore, the gradient has to be calculated with finite differences or gradient-free methods must be used. Both approaches are very costly with respect to computing time. Also, the parallelization cannot be done uniformly because the three components – mesh generator, solver, optimizer – normally use different parallelization strategies. Nevertheless, this approach can be used for some simple problems, but it is not applicable to real-world problems such as blood pumps.

We are solving the shape optimization for fluids problem within one program. We already demonstrated that this is possible for steady[4] as well as for unsteady[5] flows. In these cases we were able to show an influence of the constitutive model with respect to the optimal solution. This result showed clearly the demand for appropriate blood constitutive models as well as blood damage models. We have coupled an optimization driver with XNS, using a line search algorithm with a BFGS update method as the optimization strategy. The same approach was then used without significant changes to perform parallel optimization.

Our current work is focussed on the complex geometry deformation and new blood damage models which are more amenable to gradient computation.

Acknowledgments

The authors gratefully acknowledge the computing resources provided by the RWTH Aachen University Computing and Communication Centre, and by the John von Neumann Institute for Computing. The authors also wish to acknowledge the support the MicroMed company and of the German Science Foundation under programs GSC 111, EXC 128, SFB 540 and 401, and SPP 1253 and 1273.

References

1. D. Arora, M. Behr and M. Pasquali, *Hemolysis Estimation in a Centrifugal Blood Pump Using a Tensor-based Measure*, Artificial Organs, Vol.30 (2006): 539–547.
2. W. Frings, M. A. Hermanns, B. Mohr and B. Orth, *Report on the Jülich Blue Gene/L Scaling Workshop 2006*, Technical Report FZJ-ZAM-IB-2007-02, February 2007.
3. M. Geimer, F. Wolf, B. J. N. Wylie and B. Mohr, *Scalable Parallel Trace-Based Performance Analysis*, inSIDE Vol.4, No.2, 2006.
4. F. Abraham, M. Behr and M. Heinkenschloss, *Shape Optimization in Steady Blood Flow: A Numerical Study of Non-Newtonian Effects*, Computer Methods in Biomechanics and Biomedical Engineering, Vol.8 (2005), 127–137.
5. F. Abraham, M. Behr and M. Heinkenschloss, *Shape Optimization in Unsteady Blood Flow: A Numerical Study of Non-Newtonian Effects*, Computer Methods in Biomechanics and Biomedical Engineering, Vol.8 (2005), 201–212.

Massively Parallel Simulations of Lagrangian Plasma Turbulence

H. Homann, T. Hater, C. Beetz, C. Schwarz, J. Dreher, and R. Grauer

Institute for Theoretical Physics I
Ruhr-University Bochum, Germany
E-mail: grauer@tp1.rub.de

Some of the outstanding problems in space- and astrophysical plasma systems include hydro- or magnetohydrodynamic (MHD) turbulence (e.g. in the interstellar medium). In many situations the region under consideration contains particles advected by the turbulent flow. The statistics of these particles (Lagrangian statistics) such as scaling behaviour or mass density distributions is an active field of research. Numerical simulations of fully developed turbulence demand for high resolution, an efficient parallel implementation and a powerful computational infrastructure. The main focus of this work lies on flows which can be considered as incompressible. Nevertheless, some comparisons to compressible turbulence will be drawn. We will describe the physics behind these problems and present the numerical frameworks for solving them on massive parallel computers, such as the IBM Blue-Gene architecture.

1 Introduction

In this paper we will present numerical simulations of turbulence in plasmas and neutral flows. The focus hereby is on the computation of the Lagrangian statistics, i.e. statistical properties of particles suspended in an incompressible turbulent flow.

Turbulence is one of the most important unsolved problems in classical physics. Coherent non-linear structures in turbulent flows, such as vortex filaments (see left part in Fig. 4) and current sheets play a crucial role for the subtle statistical properties. Understanding the statistics of particles advected by a turbulent flow is a challenging theoretical and numerical problem. Simulations of fully developed turbulence make great demands on the computational infrastructure and numerical methods, especially in the case of plasma turbulence.

The incompressible simulations are carried out using a pseudo-spectral solver, called LATU. This solver will be presented in detail in the following section and demands on the computational infrastructure will be discussed. Results obtained from high-resolution numerical simulations using the LATU solver will be presented in section 4. Although, most of the research on turbulence deals with incompressible flows, one encounters compressible flows in many physical circumstances. Simulations of compressible turbulence, which develop shocks (see right part in Fig. 4) need a completely different numerical solver. We use an adaptive mesh refinement framework for hyperbolic conservation laws, *racoon*. It will be presented in section 3 and results are summarized in section 4.

2 LATU: An Incompressible MHD Solver with Passive Particles

In this section a highly parallel pseudo-spectral solver, called LATU, will be presented. The code solves the magnetohydrodynamic (MHD) equations

$$\partial_t \vec{\omega} = \nabla \times \left[\vec{v} \times \vec{\omega} - \vec{b} \times (\nabla \times \vec{b}) \right] + \mu \Delta \vec{\omega}, \tag{1}$$

$$\partial_t \vec{b} = \nabla \times (\vec{v} \times \vec{b}) + \eta \Delta \vec{b}, \tag{2}$$

$$\nabla \cdot \vec{v} = 0, \quad \nabla \cdot \vec{b} = 0, \tag{3}$$

$\vec{v}$ denoting the velocity field, related to the vorticity by $\vec{\omega} = \nabla \times \vec{v}$, and $\vec{b}$ the magnetic field which is given in non-dimensional multiples of a reference Alfvén speed. μ, η are the kinematic viscosity, magnetic diffusivity, respectively. This set of equations regards a plasma as a conducting fluid, which is appropriate in many physical situations such as in the solar wind. The left equation of (3) establishes the condition of incompressibility.

Setting $\vec{b} = 0$ in the Equations (1) to (3) one obtains the Navier-Stokes equations, which describe the motion of a neutral fluid. These equations are commonly used for studying fully developed turbulence.

The set of equations (1)–(3) are solved using a pseudo-spectral method. The underlying equations are treated in Fourier-space, while convolutions arising from non-linear terms are computed in real-space. A Fast-Fourier-Transformation (FFT) is used to switch between these two spaces. The time scheme is a (low storage) Runge-Kutta of third order. The inter-process communication uses the Message Passing Interface (MPI). More than half of the computational time is spend on the FFTs. Therefore, an efficient implementation deserves attention. The LATU-code is able to use different implementations for the FFT-algorithm. First, it can use the MPI-parallel, portable FFTW library[1] (version 2.1.5). This library decomposes the entire domain via slab-geometry, depicted in the left part of Fig. 1. Every CPU processes an ensemble of 2D-slices. This FFT parallelizes very efficiently on mid-size computers, with up to several hundred CPUs. Apparently the number of used CPUs is limited to the number of grid points in each direction. To overcome this restriction, the LATU-code uses in addition to the FFTW the San-Diego P3DFFT library. This is a parallel interface to 1D-FFT routines. For these one dimensional FFTs one can use for example the ESSL library or the FFTW3 library. The P3DFFT permits a 2D-decomposition of the computational domain, depicted in the right part of Fig. 1. Hence, the limiting restriction concerning the number of used CPUs is weakened.

Simulations of Navier-Stokes and MHD turbulence with 1024^3 collocation points each were performed using up to 512 CPUs on the IBM p690 machine of the John von Neumann Institute in Jülich. Results of scaling tests with collocation points ranging from 512^3 to 2048 on BlueGene/L using up to 16384 processors are depicted in the left part of Fig. 2. All tests were done in the virtual node mode, i.e. both cores are used for computation. By doubling the number of processes, the performance increases by 80%, the remainder is spend on the inter-process communication. A better scalability might be achieved in the communication mode. However, the overall performance of the virtual node mode is better.

A pseudo-spectral scheme is very accurate and produces negligible numerical dissipation. This is because derivatives are computed in Fourier-space and are therefore exact. Only the need for removing aliasing errors from convolutions sums and the time scheme

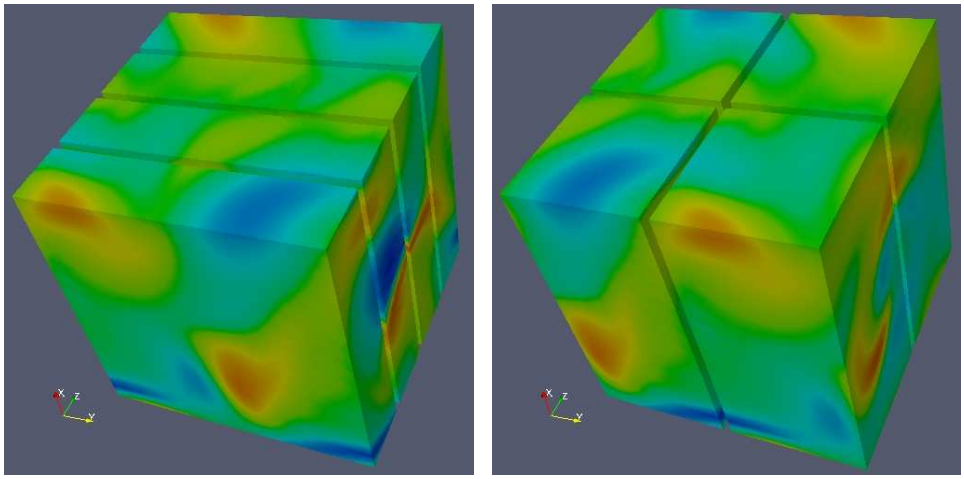


Figure 1. Partitioning of the domain for four processes. Left: slab-geometry (FFTW). Right: 2D-decomposition (P3DFFT).

yield a tiny source of dissipation. In order to choose the most adapted method of dealiasing for simulations of turbulence we compared three different strategies in the case of the Euler-equations[2]. These are obtained from the Navier-Stokes equations by neglecting viscosity and are very sensitive to the small scale accuracy of the numerical scheme. The main result is that a high-exponential cut-off is the most appropriate dealiasing strategy because it on the one hand suppresses almost all Gibbs-oscillations and on the other hand allows for significantly more physically relevant grid points than the classic 2/3-Rule.

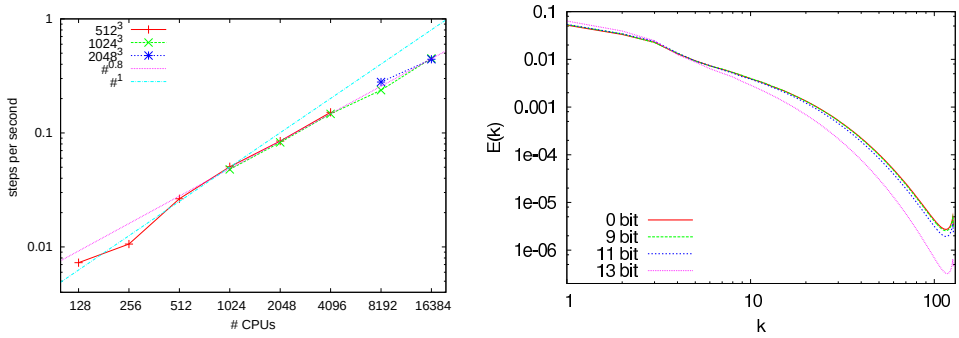


Figure 2. Left: Mixture of strong and weak scaling of the LaTu code on BlueGene/L. Right: energy-spectra for single floating-point precision simulations with an artificially reduced number of bits.

Because simulations of fully developed turbulence demand for the highest achievable computational resources, we analyzed the influence of the floating-point precision on the obtained statistical results. Therefore we artificially reduced the number of bits in the floating-point representation and compared the results to a double floating-point precision

computation. It turns out[3] that single floating-point precision is sufficient up to 4096^3 collocation points (see right part of Fig. 2) and therefore allows for a larger number of grid points by a given amount of memory. In addition many super-computers process single-precision code faster than double-precision computations.

Besides solving the Navier-Stokes and MHD equations the LATU-code is able to advect a passive scalar field according to

$$\partial_t \Theta + \vec{v} \nabla \Theta = \kappa \Delta \Theta.$$

The passive scalar Θ streams according to the velocity field and is subject to diffusion with a constant parameter κ.

An important subject in the field of turbulence is the Lagrangian statistics. This statistics deals with quantities recorded along a trajectory of a fluid element (tracer particle) exposed to a turbulent flow. In order to sample the numerical domain homogeneously and to obtain reliable statistical result within a few large-eddy turn-over times, millions of particles have to be integrated. The LATU-code is able to advect tracers and particles lighter or heavier than the surrounding fluid. This provides the possibility to analyze the influence of inertia on certain statistical quantities with a single run. All particles are treated in a parallel way using MPI. We performed simulations with up to 10^7 particles on the IBM p690 machine.

The crucial point is the interpolation scheme needed in order to obtain the velocity field at the particle positions from the numerical grid. The code uses a tri-cubic interpolation scheme which on the one hand provides a high degree of accuracy[3] and on the other hand parallelizes efficiently.

In conclusion, the LATU-code provides the opportunity to gather a lot of important quantities with a single parallel high-resolution numerical simulation of fully developed turbulence both for neutral as for conducting flows. As state of art simulations demand for resolutions of 2048^3 and millions of advected particles in order to shed light on the subtle non-linear features of self-organization of turbulence, the Blue-Gene architecture provided by the John von Neumann Institute is a key infrastructure for such simulations.

3 The Framework *racoon*

All simulations of compressible turbulence use finite volume methods and are performed using the framework *racoon*[4]. Instead of taking into account the equation, $\nabla \cdot v = 0$ (in the incompressible case), in addition to the momentum equation one now solves

$$\partial_t \rho + \nabla \cdot \rho \vec{v} = 0.$$

The pressure p is given by an isothermal equation of state.

racoon is a computational framework for the parallel, mesh-adaptive solution[5] of systems of hyperbolic conservation laws like the time-dependent Euler equations in compressible gas dynamics or Magneto-Hydrodynamics (MHD) and similar models in plasma physics. Local mesh refinement is realized by the recursive bisection of grid blocks along each spatial dimension, implemented numerical schemes include standard finite-differences as well as shock-capturing central schemes[6], both in connection with Runge-Kutta type integrators. Parallel execution is achieved through a configurable hybrid of

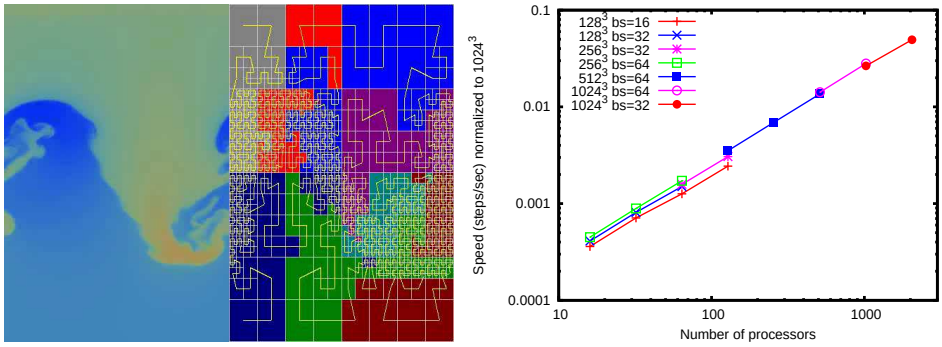


Figure 3. Left: Simulation of a Rayleigh-Taylor instability. Load balancing is based on a Hilbert curve distribution, Right: Strong scaling for different block sizes (bs) varying from 16^3 to 64^3.

POSIX-multithreading and MPI distribution with dynamic load balancing based on space-filling Hilbert curves[7] (see left part of Fig. 3).

Incompressible flows are conveniently solved using pseudo-spectral codes. Neutral incompressible flows develop coherent vortex filaments shown in the left part of Fig. 4. The formation of strong shocks in compressible gas dynamics (see right part of Fig. 4) needs a shock capturing central scheme and allows the application of adaptive mesh refining techniques (AMR)[9], so this problem is predestined to the *racoon* framework (see below). With this we examined isothermal Euler turbulence up to an effective resolution of 1024^3 cells. The right part of Fig. 4 shows the vorticity of a high Mach number compressible simulation.

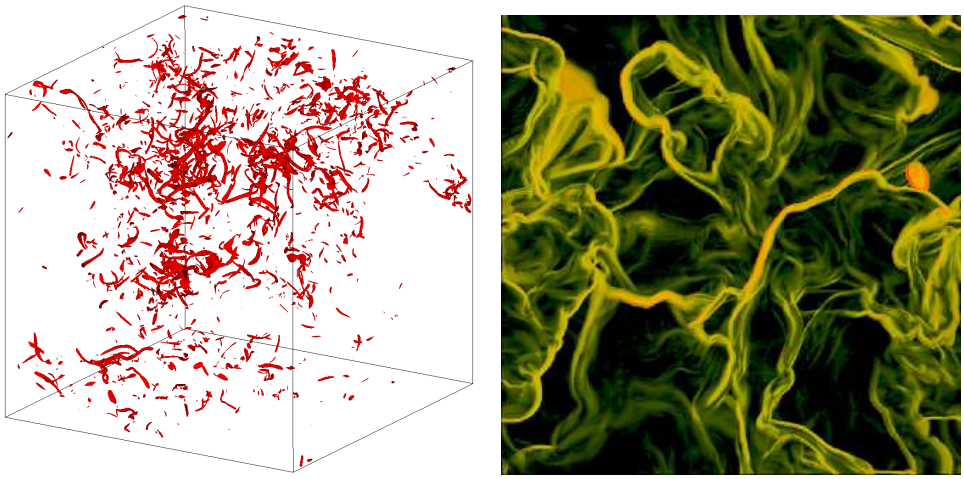


Figure 4. Dissipative structures in turbulent flows Left: Isosurface of vorticity showing vortex filaments. Right: Volume rendering of vorticity of compressible turbulence showing shock-like structures.

In compressible MHD simulation an additional problem is to assure $\nabla \cdot \vec{b} = 0$. A proper way to do this seems to be the constraint transport method[10] on a staggered grid in combination with divergence-free reconstruction[11] for the AMR.

racoon also has the ability to advect tracer particles with the flow using the same parallelization strategy as for the blocks. The main numerical work for the particle integration is spent in the interpolation routines from cell values to the actual particle positions.

Benchmarks on IBM p690 machines with 32 CPUs show that the hybrid concept in fact results in performance gain over a pure MPI parallelization, which, however requires a careful optimization of the multi-threaded implementation.

On very massively parallel machines such as the IBM BlueGene, a MPI only version is used. Scaling tests on the BlueGene/L JUBL at the FZ Jülich reveal linear scaling up to 2048 processors (see left part of Fig. 3).

4 Turbulence

Turbulence is an important and wide spread matter in today's research. From gas in molecular clouds to blood streaming through a heart valve one has to deal with turbulent flows and their properties. Although their generation is based on different forces and although they are enclosed by specific boundaries, there are features which all turbulent flows have in common.

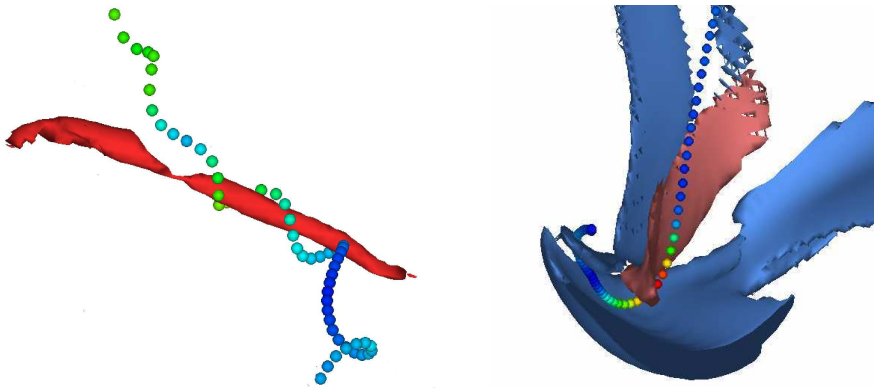


Figure 5. Particle trajectories near singular events (left: Navier-Stokes, right: MHD)

The forces and boundaries naturally act on the large scales of the motion. From these scales turbulence generates a whole range of structures of different sizes down to the smallest scales where the dissipation transforms the kinetic energy into heat. The universality sets in at scales smaller than the boundary or forcing scale down to scales larger than the dissipation scales. Here the information of the geometry and the specific dissipation mechanism of the flow is lost and the motion is completely determined by the non-linear inertial interaction of the eddies. This range is called inertial range. Physical theories often deal with fundamental features such as scaling behaviour and intermittency of this range

of scales. In order to analyze its properties numerically it is necessary to provide a large amount of scales. Numerical simulations of Euler- and Navier-Stokes-turbulence revealed that a resolution of at least 512^3 collocation points is needed to obtain an inertial range of scales within a turbulent flow.

Lagrangian statistics of turbulent neutral and magnetohydrodynamic flows has undergone a rapid development in the last 6 years to enormous progress in experimental techniques measuring particle trajectories[12]. Lagrangian statistics is not only interesting for obtaining a deeper understanding of the influence of typical coherent or nearly-singular structures in the flow but also of fundamental importance for understanding mixing, clustering and diffusion properties of turbulent astrophysical fluid and plasma flows.

Concerning the incompressible case we computed the Lagrangian statistics of MHD- and neutral turbulence. The comparison revealed the intriguing and differing influence of the flow structures on the Eulerian and Lagrangian statistics[13] (see Fig. 5). The issue of intermittency was addressed by the computation of probability density functions (PDFs) and structure functions and comparison to a multifractal model[14].

An up to now poorly understood problem in the measured Lagrangian statistics is the absence of a clear scaling range. Even the largest numerical simulations, which do show a spatial scaling range in the Eulerian framework, lack of a temporal scaling range. We addressed this issue by introducing a new time increment[15], which shows a clear scaling range (see left part in Fig. 6) and identified the large scale sweeping as a possible cause of the spoiled scaling range of the standard increment.

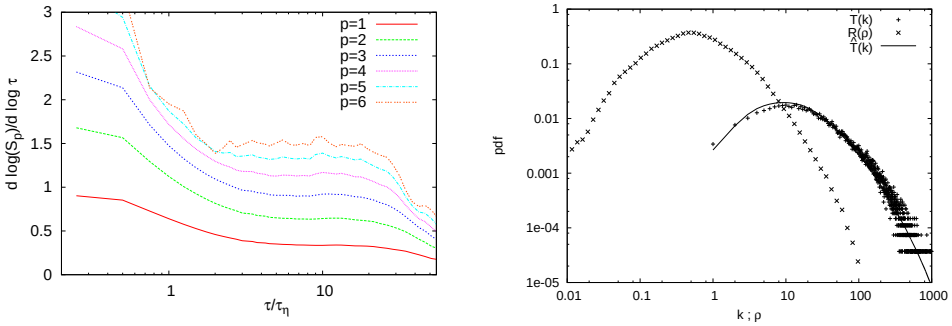


Figure 6. Left: Logarithmic derivative of structure functions of the novel time increment showing plateau and therefore a clear scaling range in incompressible turbulence, Right: The mass density PDF $R(\rho)$ (x), the measured particle distribution T (+) and the theoretical prediction $\hat{T}$ (solid line) in compressible turbulence.

A significant difference in the Lagrangian statistics between the incompressible and compressible case is found for the mass distribution of the tracers. While one finds Poisson-distributions resembling homogeneity for incompressible flows, tracers cluster in compressible flows. We found an explanation[16] for the PDF of the spatial particle distribution as a counterpart of the mass density field (see right part in Fig. 6).

References

1. FFTW library, `http://www.fftw.org`.
2. T. Grafke and H. Homann and J. Dreher and R. Grauer, *Numerical simulations of possible finite time singularities in the incompressible Euler equations: comparison of numerical methods*, to appear in Physica D, doi:10.1016/j.physd.2007.11.006.
3. H. Homann and R. Grauer, *Impact of the floating-point precision and interpolation scheme on the results of DNS of turbulence by pseudo-spectral codes*, Comp. Phys. Comm. **177**, 560–565 (2007).
4. J. Dreher and R. Grauer, *Racoon: A parallel mesh-adaptive framework for hyperbolic conservation laws*, Parallel Comp. **31**, 913–932 (2005).
5. M.J. Berger, P. Colella, *Local adaptive mesh refinement for shock hydrodynamics*, J. Comp. Phys. **82**, 64–84 (1989).
6. A. Kurganov, E. Tadmor, *New high-resolution central schemes for nonlinear conservation laws and convectiondiffusion equations*, J. Comp. Phys. **160**, 241–282 (2000).
7. G.W. Zumbusch, *On the quality of space-filling curve induced partitions*, Z. Angew. Math. Mech. **81**, 25–28 (2001).
8. R. Rabenseifer, *Hybrid Parallel Programming on HPC Platforms*, Fifth European Workshop on OpenMP, Aachen (2003).
9. A. Kritsuk, M. L. Norman, P. Padoan, *Adaptive mesh refinement for supersonic molecular cloud turbulence*, The Astrophysical Journal **638**, L25-L28 (2006).
10. U. Ziegler, *A central-constrained transport scheme for ideal magnetohydrodynamics*, J. Computat. Phys. **196**, 393–416 (2004).
11. D. S. Balsara, *Second-Order-accurate Schemes for Magnetohydrodynamics with Divergence-free Reconstruction*, The Astrophysical Journal Supplement Series **151**, 149–184 (2004).
12. A. La Porta, G. A. Voth, A. M. Crawford, J. Alexander and E. Bodenschatz, *Fluid particle accelerations in fully developed turbulence*, Nature **409**, 1017–1019 (2001).
13. H. Homann, R. Grauer, A. Busse and W.C. Müller, *Lagrangian Statistics of Navier-Stokes- and MHD-Turbulence*, J. Plasma Phys. **73**, 821–830.
14. L. Biferale, G. Boffetta, A. Celani, B. J. Devenish, A. Lanotte and F. Toschi, *Multifractal Statistics of Lagrangian Velocity and Acceleration in Turbulence*, Phys. Rev. Lett. **93**, 064502 (2004).
15. R. Friedrich, R. Grauer, H. Homann and O. Kamps, *A Corrsin-type approximation for Lagrangian Fluid Turbulence*, submitted to Phys. Rev. Lett. (2007), arXiv:0705.3132v1 [physics.flu-dyn].
16. Ch. Beetz, Ch. Schwarz, J. Dreher, and R. Grauer, *Density-PDFs and Lagrangian Statistics of highly compressible Turbulence*, submitted to Physics Letters A, arXiv:0707.1798v1 [physics.flu-dyn].

The Fine-Scale Structure of Turbulence

Jörg Schumacher[1], **Mohammad S. Emran**[1] **and Bruno Eckhardt**[2]

[1] Department of Mechanical Engineering, Technische Universität Ilmenau, 98684 Ilmenau

[2] Fachbereich Physik, Philipps-Universität Marburg, 35032 Marburg
E-mail: Joerg.Schumacher@tu-ilmenau.de

The fine-scale structure of turbulence is studied in three systems: fluid turbulence, scalar mixing, and turbulent convection. It is always characterized by large fluctuations of the corresponding field gradients which are probed by dissipation rate fields. We discuss the statistics of these quantities and the associated characteristic scales of the most intense dissipation fluctuations.

1 Introduction

Turbulence remains an outstanding fundamental problem of classical physics because of its disordered, unsteady and nonlinear dynamics which covers a wide range of scales in length and time. These variations also give rise to large fluctuations of steep gradients, which appear preferentially at the smallest scales of the flow. The amplitudes of the gradients do not follow a Gaussian distribution, and extremal values frequently exceed the mean values by orders of magnitude. This phenomenon is known as *small-scale intermittency*[1–3]. It is believed that these intense fine-scale fluctuations are intimately connected with the nonlinear cascade-like transfer of energy through the hierarchy of turbulent structures. In order to arrive at a better understanding of turbulence as a whole, we here present results from a high resolution study of the intermittent dynamics at the small-scale end of the cascade range. Such analysis is also important for the mixing in reacting and non-reacting flows, especially when a significant fraction of stirring of the concentration field takes place at the small scales of the flow. This is the case when the scalar diffusivity is small in comparison to the viscosity of the fluid.

In the absence of an analytical theory, direct numerical simulations (DNS) of the underlying fluid transport equations are the most promising approach to uncover these statistical and structural properties of turbulence, where the steepest gradients and their statistics can be resolved. Experiments cannot yet reach the finest scale filaments, despite significant progress in measurement techniques.

A fundamental quantity which probes the gradient magnitudes and displays the small-scale intermittency is the dissipation rate. In the following, we study the statistics and geometry of dissipation rate fields in three different turbulent flows: fluid turbulence, passive scalar mixing and turbulent Rayleigh-Bénard convection. For fluid turbulence, we take the energy dissipation rate of the velocity field $\mathbf{u}(\mathbf{x},t)$, given by

$$\epsilon(\mathbf{x},t) = \frac{\nu}{2}(\partial_i u_j + \partial_j u_i)^2 \,, \tag{1}$$

where ν is the kinematic viscosity of the fluid and $i,j = x,y,z$. Our second example is the *passively* advected concentration field $\theta(\mathbf{x},t)$, with dissipation rate

$$\epsilon_\theta(\mathbf{x},t) = \kappa(\partial_i \theta)^2 \,, \tag{2}$$

where κ is the diffusivity. Our final example is thermal convection, where the temperature field $T(\mathbf{x}, t)$ represents an *active* scalar field that can react back onto the flow by buoyancy. The corresponding thermal dissipation rate is given by

$$\epsilon_T(\mathbf{x}, t) = \kappa(\partial_i T)^2 \,. \tag{3}$$

with κ the thermal diffusion constant.

As we will see, extremal values of the dissipation field occur preferentially in form of thin sheets. A characteristic scale which is associated with the dissipation maxima is therefore the cross-section thickness scale of the sheets. The fluctuating nature of the dissipation generates a whole range of these scales, thus replacing the mean dissipation scale usually considered by a distribution of dissipation scales.

2 Fluid Turbulence

The continuity and Navier-Stokes equations for an incompressible fluid are given by

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{4}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{f} \,. \tag{5}$$

Here, $\mathbf{u}$ is the three-dimensional turbulent velocity field, $\mathbf{f}$ is a forcing field which sustains the statistical stationarity of turbulence and p is the (kinematic) pressure field. The dimensionless parameter that characterizes the flow is the Reynolds number based on the velocity fluctuations,

$$Re = \frac{\langle u^2 \rangle^{1/2} L}{\nu} \,. \tag{6}$$

The equations are solved by a pseudospectral method in a periodic box of side length $L = 2\pi$. More details are given in Schumacher *et al.*[4]. In contrast to many other large scale simulations with a similar number of nodes, the highest Reynolds numbers studied here are moderate: since we are interested in the small scales we cannot sacrifice small scale resolution in order to achieve high Reynolds numbers, as often done (some of the problems with such an approach are illustrated using shell models in Ref. 5). But the simulations do set a record in that they reach down to the smallest scales ever resolved.

Kolmogorov refined his classical hypothesis by incorporating the fluctuating nature of the dissipation rate field and predicted a log-normal statistics for ϵ (see Refs. 1 and 2). This is the point of reference for studies of small-scale intermittency in turbulence. Fig. 1 shows a snapshot of the dissipation field. The intermittent nature of the quantity is reflected in the few localized very-large amplitude regions which are embedded in an ambient and less strongly varying dissipation background. The extreme events will be found in the tails of the probability density function of the dissipation field. In Fig. 2, we show the PDFs of the dissipation field obtained from our high-resolution simulations for different Reynolds numbers. We see that systematic deviations of the statistics from log-normality appear in the far tails, i.e. for extremely small and large amplitudes. In other words, the log-normal model will differ for higher-order moments of the dissipation rate field. The dependence of these moments on the Reynolds number Re was derived recently by Yakhot and Sreenivasan[6]. The approach starts from the Navier-Stokes equations and

gives a relation between moments of the energy dissipation and the scaling of the structure functions in the inertial range:

$$\langle \epsilon^n \rangle \sim Re^{n+\frac{\zeta_{4n}}{\zeta_{4n}-\zeta_{4n+1}-1}} . \tag{7}$$

where ζ_n are the scaling exponents of the velocity structure function in the inertial cascade range of turbulence. Relation (7) connects the physics at small scales –the scales where the dissipation field fluctuates– with the velocity statistics in the inertial range. Such relations are also known as bridge relations. The strong resolution requirements are reflected in the large order of ζ_n. A dissipation moment of order 4 is not converged before the analytic range of the velocity structure function of order 16 is resolved! This justifies the resolution and the numerical efforts taken here. Table 1 compares our simulation results with the classical theory and Ref. 6. The recent approach agrees quite well with the simulation results while the log-normal model deviates as already seen in Fig. 2. A further consequence of Ref. 6 is that a whole range of characteristic dissipation scales exists instead of one mean scale as in the classical theory of turbulence[7].

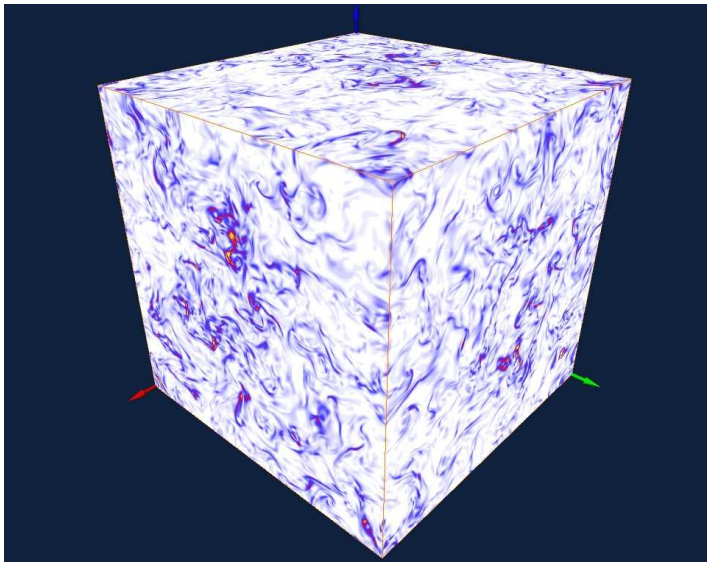


Figure 1. Contour plots of the energy dissipation rate field ϵ taken at the sideplanes of the cube. Data are from a simulation with a resolution of $N^3 = 2048^3$ grid points for a volume $V = (2\pi)^3$.

3 Passive Scalar Turbulence

In passive scalar turbulence, the concentration field $\theta(\mathbf{x}, t)$ is advected by a turbulent flow according to the advection-diffusion equation,

$$\frac{\partial \theta}{\partial t} + (\mathbf{u} \cdot \nabla)\theta = \kappa \nabla^2 \theta + f_\theta . \tag{8}$$

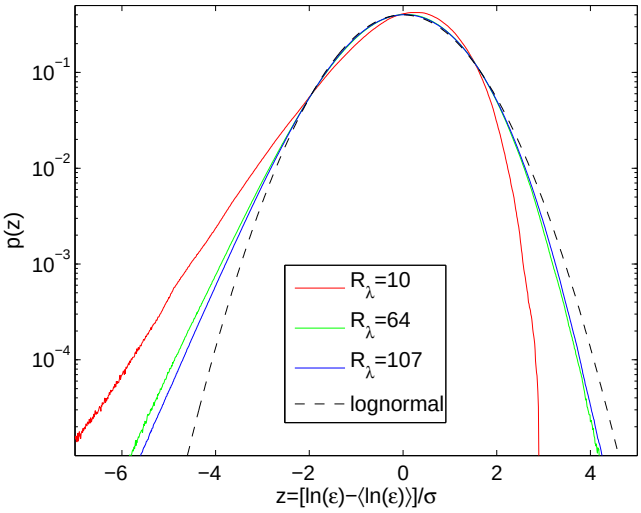


Figure 2. Probability density functions of the energy dissipation rate field at different Taylor microscale Reynolds numbers. The data are compared to the log-normal distribution which was suggested by Kolmogorov and Oboukhov in 1962. Systematic deviations from this small-scale intermittency model can be seen in particular for the rare low- and large-amplitude events which are found in the tails of the distribution.

	Theory (Yakhot)	Simulation	Kolmogorov (1962)
d_2	0.157	0.152	0.173
d_3	0.489	0.476 ± 0.009	0.465
d_4	0.944	0.978 ± 0.034	0.844

Table 1. Comparison of scaling exponents for energy dissipation rate moments of order 2, 3 and 4 with theories. In the simulations the moments and their dependence on the Reynolds number can be directly measured. Exponents are given by $d_n = n + \zeta_{4n}/(\zeta_{4n} - \zeta_{4n+1} - 1)$ in Ref. 6. and $d_n = 3(n - \zeta_{3n})/4$ in Refs. 1 and 2.

Here, f_θ is a force that sustains statistically stationary scalar fluctuations. Boundary conditions, numerical scheme and geometry are the same as for the pure fluid case. The dimensionless parameter of interest is the Schmidt number. It is given by

$$Sc = \frac{\nu}{\kappa}\,. \tag{9}$$

For $Sc > 1$, a significant fraction of the scalar is advected by the small-scale velocity field, very similar to chaotic advection[8,9]. The resulting scalar dissipation field is shown in Fig. 3. The local maxima of the dissipation field are curved sheets which result in the red coloured filaments for a plane cut as in the left panel[10]. The local cross-section thickness of these filaments was detected by a fast multiscale image segmentation algorithm[5]. Its application to the simulation data is demonstrated in the right panel. The local thickness is given then by the second principal component value which is calculated for each of the

differently coloured subfilaments. The distribution of the scales l_d is shown in Fig. 4 for different Reynolds and Schmidt numbers. Rescaling of the distributions by the corresponding mean dissipation scale (Kolmogorov scale η for the Reynolds number dependence and the Batchelor scale $\eta_B = \eta/\sqrt{Sc}$ for the Schmidt number dependence at fixed Reynolds number) results in a collapse of the data for different parameters. As in the case of fluid turbulence, we can find a whole range of *local* dissipation scales which are here directly determined from the largest amplitudes of the dissipation field.

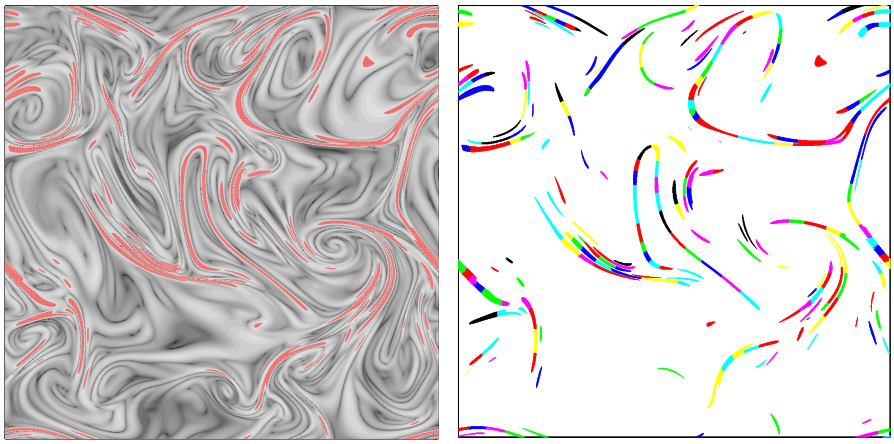


Figure 3. Left: Contour plot of a two-dimensional slice cut through the instantaneous three-dimensional scalar dissipation rate field ϵ_θ. Local maxima are replotted in red. The Schmidt number is 32 and the Taylor microscale Reynolds number is 24. The mean Batchelor scale is resolved with 2 grid cells in the simulation. Right: Reconstruction of the red coloured filaments as shown to the left by means of the fast multiscale clustering algorithm. Long filaments are composed of several subfilaments that are coloured differently.

4 Turbulent Rayleigh-Bénard Convection

In our final example we study turbulent convection, and take the influences from confining boundaries into account. We describe the temperature field within the Boussinesq approximation, and assume that the mass density of the fluid varies linearly with temperature. Temperature fluctuations drive the fluid motion by a buoyancy term that has to be substituted in the Navier-Stokes equations (5). It is given by

$$\mathbf{f}(\mathbf{x},t) = g\alpha T(\mathbf{x},t)\mathbf{e}_z\,. \tag{10}$$

α is the thermal expansion coefficient and g the gravity acceleration. The equation for the temperature field is the same as in the passive scalar case. It is given by

$$\frac{\partial T}{\partial t} + (\mathbf{u}\cdot\nabla)T = \kappa\nabla^2 T\,. \tag{11}$$

The Boussinesq equations are solved in a cylindrical cell with solid walls by a finite difference scheme. No-slip boundary conditions hold for the velocity field. The hot bottom

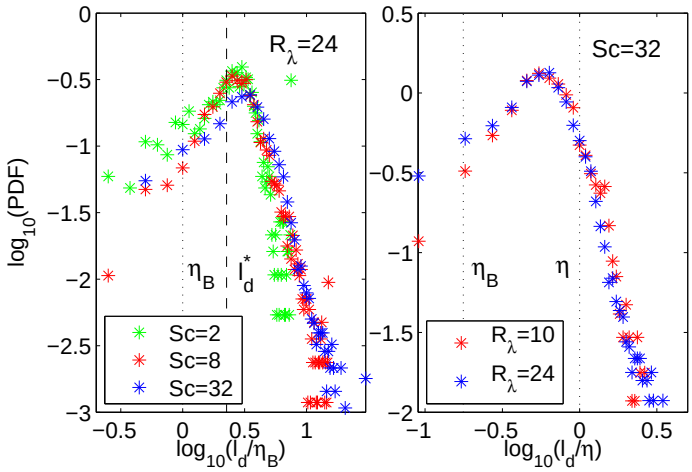


Figure 4. Distribution of the local cross section thickness l_d of the scalar dissipation rate filaments as highlighted in Fig. 3. Left panel: PDF of local dissipation scales for three different Schmidt numbers at a fixed Reynolds number. The dashed line corresponds with the theoretical value of the most probable thickness scale. Right panel: PDF of the thickness scales for two different Reynolds numbers at a fixed Schmidt number.

plate and the cold top plate are held at fixed temperatures. The side walls are adiabatic. The dimensionsless parameters are the Prandtl number Pr (equivalent to Sc), the Rayleigh number Ra which measures the magnitude of the temperature drop between top and bottom plate in z direction and the aspect ratio Γ which relates diameter D and height H

$$Pr = \frac{\nu}{\kappa}\,, \quad Ra = \frac{\alpha g H^3 \Delta T}{\nu\kappa}\,, \quad \Gamma = \frac{D}{H}\,. \tag{12}$$

Fig. 5 shows instantaneous snapshots of the total thermal dissipation rate ϵ_T (bottom) and the corresponding temperature T (top). A typical feature of turbulent convection are so-called *thermal plumes* which consist of hot blobs detaching from the boundary layer and rising into the cell. They are connected with large local thermal dissipation rate amplitudes which give the main contribution to the turbulent heat transport through the convection cell.

The temperature field is decomposed into a z-dependent mean profile $\langle T\rangle_A(z)$ and the fluctuations about the mean $\theta(\mathbf{x}, t)$. Consequently, the thermal dissipation rate can be decomposed into two contributions,

$$\langle\epsilon_T\rangle_A(z) = \epsilon_{\langle T\rangle}(z) + \langle\epsilon_\theta\rangle_A(z)\,, \tag{13}$$

where $\langle\cdot\rangle_A$ is a time-plane average at fixed $z \in [0, H]$. Fig. 6 shows the vertical profiles of the three different terms in (13). It can be seen that the dominant part of the thermal dissipation close to the plates comes from the mean profile term. This suggests that the plumes affect the cell to a certain height only, an open question of present research on the subject. Towards the centre of the cell the total magnitude of the dissipation drops and is mainly due to the thermal fluctuations. Our present studies indicate that the statistical properties of the thermal dissipation rate field in the centre of the cell are very similar to

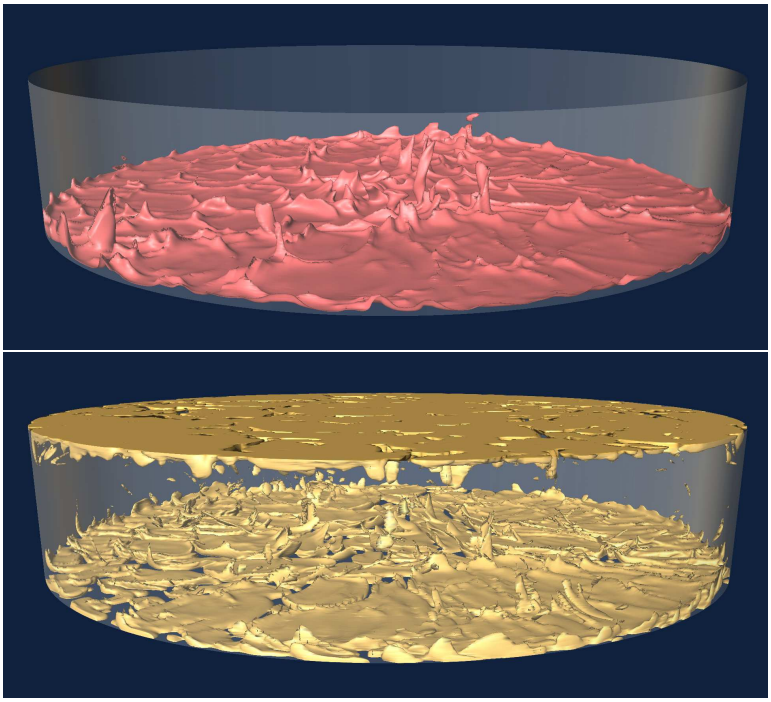


Figure 5. Snapshots of the temperature field T (top) and the corresponding thermal dissipation rate field ϵ_T (bottom) for turbulent convection in a cell with aspect ratio $\Gamma = 5$ at a Rayleigh number $Ra = 10^7$. The temperature level $T = 0.75T_{bottom}$ is shown. The isolevel of the thermal dissipation rate field is 0.02.

those of the passive scalar dissipation rate[9]. In both cases we find again deviations of the PDF from log-normality as shown for the velocity field in Fig. 2. The detailed study of the statistics of the thermal dissipation rate in the boundary layers is currently in progress.

Acknowledgments

We would like to thank W. J. A. Dahm, P. E. Hamlington, C. R. Doering, K. R. Sreenivasan, R. Verzicco and V. Yakhot for fruitful discussions. The research is supported by the Deutsche Forschungsgemeinschaft, by the German Academic Exchange Service (DAAD) and the DEISA consortium. We would like to thank the John von Neumann Institute for Computing for the steady support of our research. JS thanks especially M. Boltes, S. Habbinga, M.-A. Hermanns, M. Rambadt and H. Zilken for their support during the phase of the DEISA project "eta2006" and M. Pütz (IBM Deutschland) for getting started at the Blue Gene/L JUBL.

References

1. A. N. Kolmogorov, *A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number*, J. Fluid Mech. **13**, 82-85, 1962.

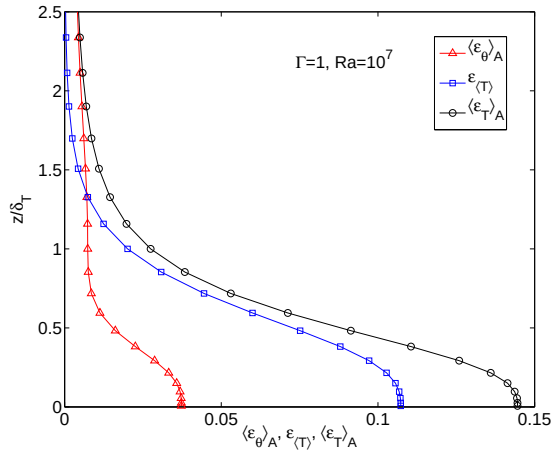


Figure 6. Vertical height dependence of the two different contributions to the total thermal dissipation rate ϵ_T as given by (13). Data are for a Rayleigh number $Ra = 10^7$ and an aspect ratio $\Gamma = 5$. The numerical grid is $N_\phi \times N_r \times N_z = 361 \times 257 \times 128$.

2. A. M. Oboukhov, *Some specific features of atmospheric turbulence*, J. Fluid Mech. **13**, 77-81, 1962.
3. K. R. Sreenivasan and R. A. Antonia, *The phenomenology of small-scale turbulence*, Annu. Rev. Fluid Mech. **29**, 435-472, 1997.
4. J. Schumacher, K. R. Sreenivasan and V. Yakhot, *Asymptotic scaling exponents from low-Reynolds number flows*, New J. Phys. **9**, 89, 2007.
5. J. C. Bowman, C. R. Doering, B. Eckhardt, J. Davoudi, M. Roberts and J. Schumacher, *Links between dissipation, intermittency, and helicity in the GOY model revisited*, Physica D **218**, 1-10, 2007.
6. V. Yakhot and K. R. Sreenivasan, *Towards a dynamical theory of multifractals in turbulence*, Physica A **343**, 147-155, 2004.
7. J. Schumacher, *Sub-Kolmogorov-scale fluctuations in fluid turbulence*, Europhys. Lett. **80**, 54001, 2007.
8. J. Schumacher, K. R. Sreenivasan and P. K. Yeung, *Very fine structures in scalar mixing*, J. Fluid Mech. **531**, 113-129, 2005.
9. J. Schumacher and K. R. Sreenivasan, *Statistics and geometry of passive scalar fields in turbulence*, Phys. Fluids **17**, 125107, .
10. J. Schumacher, H. Zilken, B. Eckhardt and K. R. Sreenivasan, *Scalar dissipation fronts in high-Schmidt number mixing (Gallery of Nonlinear Images)*, Chaos **15**, 041105, 2005.
11. D. Kushnir, J. Schumacher and A. Brandt, *Geometry of intensive scalar dissipation events in turbulence*, Phys. Rev. Lett. **97**, 124502, 2006.

Lyapunov Instabilities of Extended Systems

Hong-liu Yang and Günter Radons

Institute of Physics, Chemnitz University of Technology, D-09107 Chemnitz, Germany
E-mail: {hongliu.yang, radons}@physik.tu-chemnitz.de

Here we review our current results on Lyapunov spectra and Lyapunov vectors (LVs) of various extended systems with continuous symmetries. The major part of the article is devoted to the study of Lennard-Jones fluids in one- and two-dimensional spaces. By using the newly introduced LV correlation functions, we demonstrate that the LVs with $\lambda \approx 0$ are highly dominated by a few components with low wave numbers, which implies the existence of hydrodynamic Lyapunov modes in soft-potential systems. Despite the wave-like character of the LVs, no step-like structure exists in the Lyapunov spectrum of the systems studied here, in contrast to the hard-core case. Studies on dynamical LV structure factors conclude that HLMs in Lennard-Jones fluids are propagating. We also briefly outline our current results on the universal features of HLMs in a class of spatially extended systems with continuous symmetries. HLMs in Hamiltonian and dissipative systems are found to differ both in respect of spatial structure and in the dynamical evolution.

1 Introduction

One of the most successful theories in modern science is statistical mechanics, which allows us to understand the macroscopic (thermodynamic) properties of matter from a statistical analysis of the microscopic (mechanical) behaviour of the constituent particles. In spite of this, using certain probabilistic assumptions such as Boltzmann's *Stosszahlansatz* causes the lack of a firm foundation of this theory, especially for non-equilibrium statistical mechanics. Fortunately, the concept of chaotic dynamics developed in the 20th century is a good candidate for accounting for these difficulties. Instead of the probabilistic assumptions, the dynamical instability of trajectories can make available the necessary fast loss of time correlations, ergodicity, mixing and other dynamical randomness. It is generally expected that dynamical instability is at the basis of macroscopic transport phenomena and that one can find certain connections between them. Some beautiful theories in this direction were already developed in the past decade[1], where the Lyapunov exponents were related to certain transport coefficients.

Very recently, molecular dynamics simulations on hard-core systems revealed the existence of regular collective perturbations corresponding to the smallest positive Lyapunov exponents (LEs), named hydrodynamic Lyapunov modes[2]. This provides a new possibility for the connection between Lyapunov vectors, a quantity characterizing the dynamical instability of trajectories, and macroscopic transport properties. A lot of work[3–8] has been done to identify this phenomenon and to find out its origin. The appearance of these modes is commonly thought to be due to the conservation of certain quantities in the systems studied[3–7]. A natural consequence of this expectation is that the appearance of such modes might not be an exclusive feature of hard-core systems and might be generic to a large class of Hamiltonian systems. However, until very recently, these modes have only been identified in the computer simulations of hard-core systems[3,8].

In this article, we give an overview of our recent results on Lyapunov instabilities of various extended systems with continuous symmetries, especially on the identification

and characterization of hydrodynamic Lyapunov modes. The major part of the article is devoted to the investigation of Lennard-Jones fluids, wherein the HLMs are, for the first time, identified in systems with soft-potential interactions[9,10]. Our new technique, based on a spectral analysis of LVs, shows strong evidence that hydrodynamic Lyapunov modes do exist in these cases. In section 6, we will go beyond the many-particle systems and show some universal features of HLMs in a large class of extended systems. Details of these investigations can be found in our publications[11–13].

2 Numerical Method for Determining Lyapunov Exponents and Vectors

2.1 Standard Method

The equations of motion for a many-body system may always be written as a set of first order differential equations $\dot{\Gamma}(t) = F(\Gamma(t))$, where Γ is a vector in the D-dimensional phase space. The tangent space dynamics describing infinitesimal perturbations around a reference trajectory $\Gamma(t)$ is given by

$$\delta\dot{\Gamma} = M(\Gamma(t)) \cdot \delta\Gamma \tag{1}$$

with the Jacobian $M = \frac{dF}{d\Gamma}$. The time averaged expansion or contraction rates of $\delta\Gamma(t)$ are given by the Lyapunov exponents. For a $D-$dimensional dynamical system there exist in total D Lyapunov exponents for D different directions in tangent space. The orientation vectors of these directions are the Lyapunov vectors $e^{(\alpha)}(t)$, $\alpha = 1,\cdots,D$.

For the calculation of the Lyapunov exponents and vectors the offset vectors have to be reorthogonalized periodically, either by means of Gram-Schmidt orthogonalization or QR decomposition[14]. To obtain scientifically useful results, one needs large particle numbers and long integration times for the calculation of certain long time averages. This enforces the use of parallel implementations of the corresponding algorithms. It turns out that the repeated reorthogonalization is the most time consuming part of the algorithm.

2.2 Parallel Realization

As parallel reorthogonalization procedures we have realized and tested several parallel versions of Gram-Schmidt orthogonalization and of QR factorization based on blockwise Householder reflection. The parallel version of classical Gram-Schmidt (CGS) orthogonalization is enriched by a reorthogonalization test which avoids a loss of orthogonality by dynamically using iterated CGS. All parallel procedures are based on a 2-dimensional logical processor grid and a corresponding block-cyclic data distribution of the matrix of offset vectors. Row-cyclic and column-cyclic distributions are included due to parameterized block sizes, which can be chosen appropriately. Special care was also taken to offer a modular structure and the possibility for including efficient sequential basic operations, such as those from BLAS, in order to efficiently exploit the processor or node architecture.

Performance tests of parallel algorithms have been done on a Beowulf cluster, a cluster of dual Xeon nodes, and an IBM Regatta p690+. Results can be found in[15].

3 Correlation Functions for Lyapunov Vectors

In the spirit of molecular hydrodynamics[16], we introduced in[9,10] a dynamical variable called *LV fluctuation density*,

$$u^{(\alpha)}(r,t) = \sum_{j=1}^{N} \delta x_j^{(\alpha)}(t) \cdot \delta(r - r_j(t)), \tag{2}$$

where $\delta(z)$ is Dirac's delta function, $r_j(t)$ is the position coordinate of the j-th particle, and $\{\delta x_j^{(\alpha)}(t)\}$ is the coordinate part of the α-th Lyapunov vector at time t. The spatial structure of LVs is characterized by the *static LV structure factor* defined as

$$S_u^{(\alpha\alpha)}(k) = \int \langle u^{(\alpha)}(r,0) u^{(\alpha)}(0,0) \rangle e^{-jk\cdot r} dr, \tag{3}$$

which is simply the spatial power spectrum of the LV fluctuation density. Information on the dynamics of LVs can be extracted via the *dynamic LV structure factor*, which is defined as

$$S_u^{(\alpha\alpha)}(k,\omega) = \int\int \langle u^{(\alpha)}(r,t) u^{(\alpha)}(0,0) \rangle e^{-jk\cdot r} e^{j\omega t} dr dt. \tag{4}$$

With the help of these quantities the controversy[2,3] about the existence of hydrodynamic Lyapunov modes in soft-potential systems has been successfully resolved[9].

4 Numerical Results for 1d Lennard-Jones Fluids

4.1 Models

The Lennard-Jones system studied has the Hamiltonian $H = \sum_{j=1}^{N} mv_i^2/2 + \sum_{j<l} V(x_l - x_j)$. where the interaction potential among particles is $V(r) = 4\epsilon\left[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}\right] - V_c$ if $r \leq r_c$ and $V(r) = 0$ otherwise with $V_c = 4\epsilon\left[(\frac{\sigma}{r_c})^{12} - (\frac{\sigma}{r_c})^{6}\right]$. Here the potential is truncated in order to lower the computational burden.

The system is integrated using the velocity form of the Verlet algorithm with periodic boundary conditions. The standard method invented by Benettin et al. and Shimada and Nagashima[14] is used to calculate the Lyapunov characteristics of the systems studied. Other technical details can be found in Ref.[9]. Throughout this paper, the particle number is typically denoted by N, the length of the system by L and the temperature by T.

4.2 Smooth Lyapunov Spectrum with Strong Short-Time Fluctuations

The Lyapunov spectrum for the case $N = 100$, $L = 1000$ and $T = 0.2$ is shown in Fig. 1. Only half of the spectrum is shown here, since all LEs of Hamiltonian systems come in pairs according to the conjugate-pairing rule. In the enlargement shown in the inset of Fig. 1 for the part near $\lambda^{(\alpha)} \approx 0$, one can not see any step-wise structure in the Lyapunov spectrum, in contrast to the case of hard-core systems[3]. This is the typical result obtained for our soft potential system.

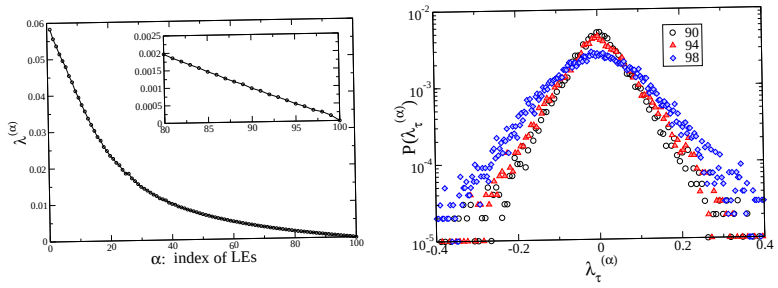


Figure 1. Left: Lyapunov spectrum. Right: Distribution of the finite-time Lyapunov exponent $\lambda_\tau^{(\alpha)}$ where τ is equal to the period of re-orthonormalization.

The fluctuations in local instabilities of trajectories is demonstrated by means of the distribution of finite-time LEs. By definition, finite-time Lyapunov exponents λ_τ measure the expansion rate of trajectory segments of the duration τ. In Fig. 1, such distributions are presented for some LEs in the regime $\lambda \approx 0$. Fluctuations of the finite time Lyapunov exponents are quite large compared to the difference between their mean values, i.e., $\sigma(\lambda_\tau^{(\alpha)}) \equiv \sqrt{\langle {\lambda_\tau^{(\alpha)}}^2 \rangle - \langle \lambda_\tau^{(\alpha)} \rangle^2} \gg |\lambda^{(\alpha)} - \lambda^{(\alpha+1)}|$. Here, $\langle \cdots \rangle$ means time average. The strong fluctuations in local instabilities constitute one of the possible reasons for the disappearance of the step-wise structures in the Lyapunov spectra. They could also cause the mixing of nearby Lyapunov vectors. The mixing may be at the basis of the intermittency observed in the time evolution of the spatial Fourier transformation of LVs (see Sect. 4.2.1).

4.2.1 Intermittency in Time Evolution of Instantaneous Static LV Structure Factors

Based on the spatial Fourier transformation of $u^{(\alpha)}(x,t)$

$$\tilde{u}_k^{(\alpha)}(t) = \int u^{(\alpha)}(x,t) \exp(-ikx) dx = \sum_{j=1}^{N} \delta x_j^{(\alpha)} \cdot \exp[-ik \cdot x_j(t)] \tag{5}$$

we introduce a quantity called *instantaneous static LV structure factor*, which reads

$$s_{uu}^{(\alpha)}(k,t) \equiv |\tilde{u}_k^{(\alpha)}(t)|^2. \tag{6}$$

It is nothing but the instantaneous spatial power spectrum of $u^{(\alpha)}(x,t)$.

The time evolution of the instantaneous static LV structure factor $s_{uu}^{(95)}(k,t)$ for Lyapunov vector No. 95 is shown in Fig. 2 as an example. Two quantities are recorded as time goes on. One is the peak wave-number k_*, which marks the position of the highest peak in the spectrum $s_{uu}^{(\alpha)}(k,t)$ (see Fig. 2). The other is the spectral entropy $H_s(t)$, which measures the distribution property of the spectrum $s_{uu}^{(\alpha)}(k,t)$. It is defined as:

$$H_s(t) = -\sum_{k_i} s_{uu}^{(\alpha)}(k_i,t) \ln s_{uu}^{(\alpha)}(k_i,t). \tag{7}$$

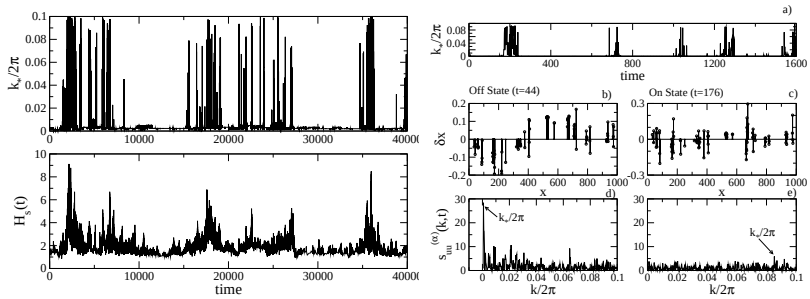


Figure 2. Left: Intermittent behaviours of the peak wave-number k_* and spectral entropy $H_s(t)$ for the spatial Fourier spectrum of $u^{(95)}(x,t)$. Right: a) Variation of the peak wave number k_* with time. b),c) Two typical snapshots of LV_{95}, *off* and *on* state at $t = 44$ and 176 respectively. d),e) their spatial Fourier transform.

A smaller value of $H_s(t)$ means that the spectrum $s_{uu}^{(\alpha)}(k,t)$ is highly concentrated on a few values of k, i.e., these components dominate the behaviour of the LV. Both of these quantities behave intermittently, as shown in Fig. 2. Large intervals of nearly constant low values (*off state*) are interrupted by short period of bursts (*on state*) where they have large values. Details of typical *on* and *off* states are shown in the right part of Fig. 2. One can see that the off state is dominated by low wave-number components (see the sharp peak at low wave-number k_*), while the on state is more noisy and there are no significant dominant components. This intermittency in the time evolution of the instantaneous static LV structure factors is a typical feature of soft potential systems. It is conjectured that this is a consequence of the mixing of nearby LVs caused by the wild fluctuations of local instabilities. Due to the mutual interaction among modes, the hydrodynamic Lyapunov modes in the soft potential systems are only of finite life-time. In the dynamic Lyapunov structure function estimated, the peak representing the propagating (or oscillating) Lyapunov modes is of finite width. This is support for our conjecture that the hydrodynamic Lyapunov modes are of finite life-time.

4.2.2 Dispersion Relation of Hydrodynamic Lyapunov Modes

Now, we consider the static LV structure factor $S_{uu}^{(\alpha)}(k)$, which is the long-time average of the instantaneous quantity $s_{uu}^{(\alpha)}(k)$. Two cases with $L = 1000$ and 2000 are shown in Fig. 3. It is not hard to recognize the sharp peak at $\lambda \approx 0$ in the contour plot of the spectrum. With increasing Lyapunov exponents, the peak shifts to the larger wave number side. A dashed line is plotted to make clear how the wave number of the peak k_{max} changes with $\lambda^{(\alpha)}$.

All of our results shown above provide strong evidence of the fact that the Lyapunov vectors corresponding to the smallest positive LEs in our 1d Lennard-Jones system are highly dominated by a few components with small wave numbers, i.e, they are similar to the Hydrodynamic Lyapunov modes found in hard-core systems. The wave-like character becomes weaker and weaker as the value of the LE is increased gradually from zero.

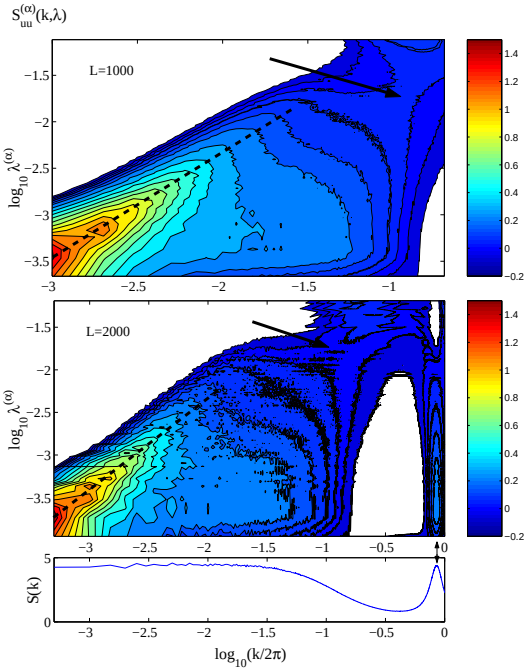


Figure 3. Contour plot of the spectra $S_{uu}^{(\alpha)}(k)$ for $L = 1000$ and 2000. A ridge structure can easily be recognized in the regime $k \approx 0$ and $\lambda \approx 0$. To guide the eyes, a dashed line is plotted to show how the peak wave-number k_{max} changes with λ.

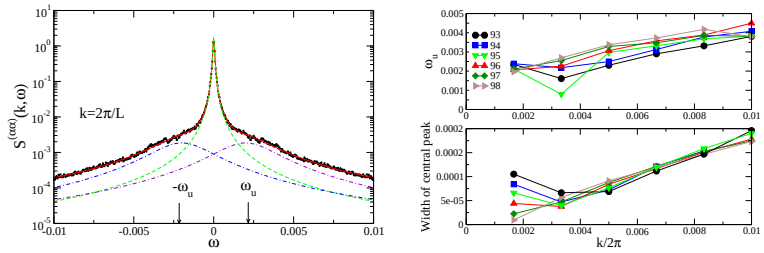


Figure 4. Left: Dynamic LV structure factor $S_u^{(\alpha\alpha)}(k,\omega)$ for $\alpha = 96$ and $k = 2\pi/L$. The full line results from a 3-pole fit. The corresponding decomposition into three Lorentzians is also shown. Right: Dispersion relations $\omega^{(\alpha)}(k)$ (top) and the $k-$dependence of the width of the central peak (bottom) obtained from 3-pole approximations.

4.3 Dynamic LV Structure Factors

More detailed information about the dynamical evolution of Lyapunov vectors can be obtained from the dynamic LV structure factors $S_u^{(\alpha\alpha)}(k,\omega)$, which encode in addition to the structural also the temporal correlations. In Fig. 4 we show a typical example for

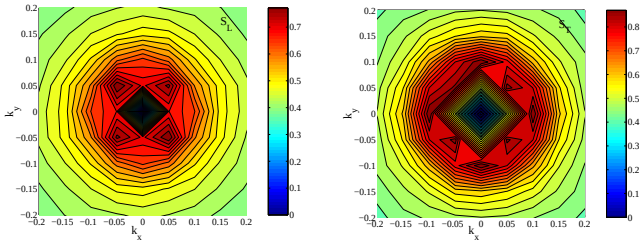


Figure 5. Contour plots of $S_L^{(\alpha\alpha)}(\vec{k})$ and $S_T^{(\alpha\alpha)}(\vec{k})$ of a LV with $\alpha = 140$ in a $2d$ system with $N = 100$, $T = 0.8$ and $L_x \times L_y = 20 \times 20$. Obviously, the longitudinal and transverse components behave differently.

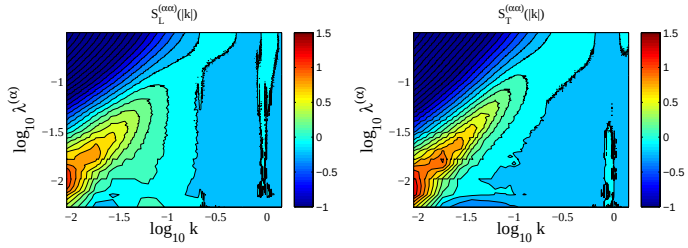


Figure 6. Contour plots of $S_L^{(\alpha\alpha)}(k)$ and $S_T^{(\alpha\alpha)}(k)$ (left) and the corresponding dispersion relation $\lambda(k)$ (right) of the hydrodynamic Lyapunov modes in a $2d$ system with $N = 100$, $T = 0.8$ and $L_x \times L_y = 5 \times 120$.

$S_u^{(\alpha\alpha)}(k,\omega)$. It consist of a central "quasi-elastic" peak with shoulders resulting from dynamical excitations quite similar to the dynamic structure factor $S(k,\omega)$ of fluids[16]. In order to extract the dynamical information we use a 3-pole approximation for $S_u^{(\alpha\alpha)}(k,\omega)$, which amounts to fitting the latter by a superposition of three Lorentzians, one central peak at $\omega = 0$ and two symmetric peaks located at $\omega = \pm\omega_u(k)$. The fits are also shown in the figure. They describe the frequency dependence of $S_u^{(\alpha\alpha)}(k,\omega)$ quite well. These fits allow us to extract the dispersion relations $\omega^{(\alpha)}(k)$ for each of the hydrodynamic Lyapunov modes with index α. The results are shown in Fig. 4 for several of the Lyapunov modes. Clearly, this tells us that a Lyapunov mode corresponding to exponent λ is characterized, apart from the dominating wave number $k(\lambda)$, by a typical frequency $\omega(k(\lambda))$. Because $\frac{d\omega}{dk}$ is non-vanishing, this implies propagating wave-like excitations. The full LV dynamics of the soft-potential system treated here, however, is more complex than that of the hard-core systems. For instance, the peaks in $S_u^{(\alpha\alpha)}(k,\omega)$ are of finite width (see Fig. 4). This fact is consistent with our observation that several quantities characterizing the dynamical aspect of Lyapunov vectors evolve erratically in time (see Sec. 4.2.1).

5 Lyapunov Modes in 2d Lennard-Jones Fluids

In isotropic fluids with $d > 1$ the static LV structure factor $S_u^{(\alpha\alpha)}(\vec{k})$ becomes a second rank tensor. Cartesian components $S_{\mu\nu}^{(\alpha\alpha)}(\vec{k})$ of $S_u^{(\alpha\alpha)}(\vec{k})$ can be expressed in terms

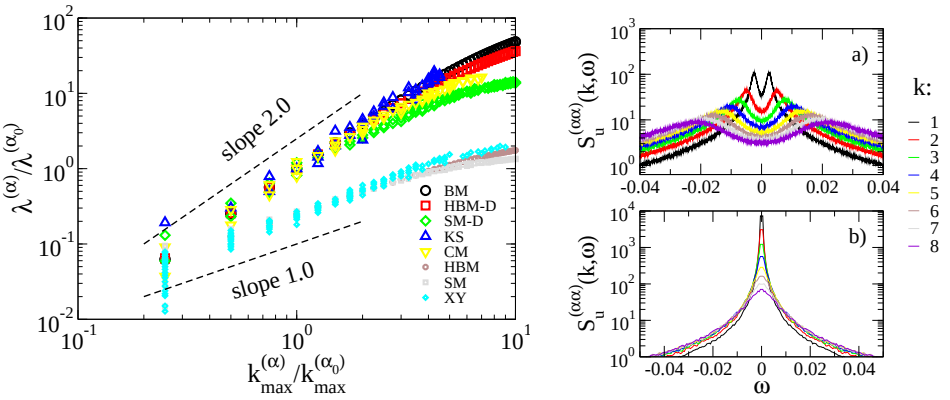


Figure 7. Left: The λ-k dispersion relations for various extended systems with continuous symmetries. The normalized data for different systems collapse on two master curves. These results strongly support our conjecture that there are two classes of systems with $\lambda \sim k$ and $\lambda \sim k^2$ respectively. Systems in the group with $\lambda \sim k$ include Eq.(8) with $f(z) = \frac{1}{2\pi}\sin(2\pi z)$ (SM), Eq.(8) with $f(z) = 2z \pmod 1$ (HBM) and the 1d XY model (XY). Systems belonging to class $\lambda \sim k^2$ are Eq.(10) with $f(z) = \frac{1}{2\pi}\sin(2\pi z)$ (CM), Eq.(10) with $f(z) = 2z \pmod 1$ (BM), Eq.(8) with $f(z) = \frac{1}{2\pi}\sin(2\pi z)$ and $\gamma = 0.7$ (SM-D), Eq.(8) with $f(z) = 2z \pmod 1$ and $\gamma = 0.7$ (HBM-D) and the 1d Kuramoto-Sivashinsky equation (KS). Right: Dynamic LV structure factors $S_u^{(\alpha\alpha)}(k,\omega)$ for a) coupled standard maps, Eq.(8) with $\epsilon = 1.3$; b) coupled circle maps, Eq.(10) with $\epsilon = 1.3$.

of longitudinal and transverse correlation functions $S_L^{(\alpha\alpha)}$ and $S_T^{(\alpha\alpha)}$ as $S_{\mu\nu}^{(\alpha\alpha)}(\vec{k}) = \hat{k}_\mu\hat{k}_\nu S_L^{(\alpha\alpha)}(k) + (\delta_{\mu\nu} - \hat{k}_\mu\hat{k}_\nu)S_T^{(\alpha\alpha)}(k)$ with $\hat{k}_\mu = (\vec{k}/k)_\mu$. As an example, we presented in Fig. 5 the contour plot of the two correlation functions S_L and S_T for LV No. 140 of a two-dimensional Lennard-Jones system with $N = 100$, $T = 0.8$ and $L_x \times L_y = 20 \times 20$. The difference between the two components is quite obvious. However, as can be seen from Fig. 6, $S_L^{(\alpha\alpha)}(k)$ and $S_T^{(\alpha\alpha)}(k)$ for two-dimensional cases behave similar to the one-dimensional case shown in Fig. 3. This fact implies the existence of hydrodynamic Lyapunov modes also in two-dimensional cases. In addition, both the longitudinal and transverse components are characterized by a linear dispersion relation, which has been found to be typical of Hamiltonian systems[12,13]. Further numerical simulations show that the transverse modes are non-propagating , in contrast to the longitudinal components.

6 Universal Features of Lyapunov Modes in Spatially Extended Systems with Continuous Symmetries

Relying on the LV correlation function method, we have up to now successfully identified the existence of HLMs in the following spatially extended systems:
Coupled map lattices (CMLs) with either Hamiltonian or dissipative local dynamics

$$v_{t+1}^l = (1-\gamma)v_t^l + \epsilon[f(u_t^{l+1} - u_t^l) - f(u_t^l - u_t^{l-1})] \tag{8}$$

$$u_{t+1}^l = u_t^l + v_{t+1}^l \tag{9}$$

and

$$u_{t+1}^l = u_t^l + \epsilon[f(u_t^{l+1} - u_t^l) - f(u_t^l - u_t^{l-1})]. \tag{10}$$

Dynamic XY model with the Hamiltonian

$$H = \sum_i \dot{\theta}_i + \epsilon \sum_{ij} [1 - cos(\theta_j - \theta_i)]. \tag{11}$$

Kuramoto-Sivashinsky equation

$$h_t = -h_{xx} - h_{xxxx} - h_x^2. \tag{12}$$

A common feature of these systems is that they all hold certain continuous symmetries and conserved quantities, which have been shown to be essential for the occurrence of Lyapunov modes[12]. Our numerical simulations and analytical calculations indicate that these systems fall into two groups with respect to the nature of hydrodynamic Lyapunov modes. To be precise, the dispersion relations are characterized by $\lambda \sim k$ and $\lambda \sim k^2$ in Hamiltonian and dissipative systems respectively, as Fig. 7 indicates. Moreover, the HLMs in Hamiltonian systems are propagating, whereas those in dissipative systems show only diffusive motion. Examples of dynamic LV structure factors for two CMLs are presented in the right row of Fig. 7. In a), each spectrum has two sharp symmetric side-peaks located at $\pm\omega_u$. Furthermore, $\omega_u \simeq \pm c_u k$ for $k \geq 2\pi/L$. These facts suggest that the HLMs in coupled standard maps are propagating. The spectrum of coupled circle maps in b) has only a single central peak and can be well approximated by a Lorentzian curve[12], which implies that the HLMs in this system fluctuate diffusively. In addition, no step structures in Lyapunov spectra have been found in contrast to the hard-core systems. The quantities characterizing the dynamical evolutions of LVs in these systems exhibit intermittent behaviour.

7 Conclusion and Discussion

We have presented numerical results for the Lyapunov instability of Lennard-Jones systems. Our simulations show that the step-wise structures found in the Lyapunov spectrum of hard-core systems disappear completely here. This is presumed to be the result of the strong fluctuations in the finite-time LEs[3]. A new technique based on the spatial Fourier spectral analysis is employed to reveal the vague long wave-length structure hidden in LVs. In the resulting spatial Fourier spectrum of LVs with $\lambda \simeq 0$, a significantly sharp peak with low wave-number is found. This serves a strong evidence of the existence of hydrodynamic Lyapunov modes in soft-potential systems[18]. The disappearance of the step-structures and the survival of the hydrodynamic Lyapunov modes show that the latter are more robust and essential than the former. Studies on dynamical LV structure factors provide evidence that longitudinal HLMs in Lennard-Jones fluids are propagating. Going beyond many-particle systems, we have shown that, for a large class of extended systems, HLMs of Hamiltonian and dissipative cases are different both in respect of spatial structure and in the dynamical evolution.

Acknowledgments

We thank W. Just, W. Kob, A. Latz, A. S. Pikovsky and H. A. Posch for fruitful discussions. Special thanks go to W. Kob for providing us with the code of molecular dynamics simulations and to G. Rünger and M. Schwind for the help on the parallel algorithms. We acknowledge financial support from the DFG within SFB393 "Parallele Numerische Simulation für Physik und Kontinuumsmechanik" and Ra416/6-1 and a grant of computer time provided by the John von Neumann Institute for Computing.

References

1. J.P. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*. Cambridge University Press, Cambridge (1999).
2. H.A. Posch and R. Hirschl, *Simulation of billards and of hard-body fluids*. In: Szasz, D. (ed) Hard Ball Systems and the Lorenz Gas. Springer, Berlin (2000).
3. C. Forster, R. Hirschl, H.A. Posch and Wm.G. Hoover, Physics D, **187**, 294-310 (2004).
4. J.-P. Eckmann and O. Gat, J. Stat. Phys., **98**, 775-798 (2000).
5. S. McNamara and M. Mareschal, Phys. Rev. E, **64**, 051103 (2001).
6. A. de Wijn and H. van Beijeren, Phys. Rev. E, **70**, 016207 (2004).
7. T. Taniguchi and G.P. Morriss, Phys. Rev. E, **65**, 056202 (2002); Phys. Rev. E, **68**, 026218 (2003).
8. Wm.G. Hoover, H.A. Posch, C. Forster, C. Dellago and M. Zhou, J. Stat. Phys., **109**, 765-776 (2002).
9. H.-L. Yang and G. Radons, Phys. Rev. E, **71**, 036211 (2005), see also arXiv:nlin.CD/0404027.
10. G. Radons and H.-L. Yang, arXiv:nlin.CD/0404028.
11. H.-L. Yang and G. Radons, Phys. Rev. Lett. **96**, 074101 (2006).
12. H.-L. Yang and G. Radons, Phys. Rev. E **73**, 016202 (2006).
13. H.-L. Yang and G. Radons, Phys. Rev. E **73**, 016208 (2006).
14. G. Benettin, L. Galgani and J.M. Strelcyn, Phys. Rev. A, **14**, 2338-2345, (1976); I. Shimada and T. Nagashima, Prog. Theor. Phys., **61**, 1605-1616 (1979).
15. G. Radons, G. Rünger, M. Schwind, and H.-L Yang, In J. Dongarra, K. Madsen, and J. Wasniewski (eds.): Applied Parallel Computing, Proceedings of PARA04, Lyngby, June 20-23, 2004, Lecture Notes of Computer Science **3732**, pp. 1131, Springer, Berlin (2006).
16. J.P. Boon and S. Yip, *Molecular Hydrodynamics*. McGraw-Hill, New York (1980).
17. J.-P. Eckmann, C. Forster, H.A. Posch and E. Zabey, J. Stat. Phys., **118**, 795-811 (2005).
18. C. Forster and H.A. Posch, New J. Phys., **7**, 32 (2005), see also arXiv:nlin.CD/0409019.

Already published:

Modern Methods and Algorithms of Quantum Chemistry - Proceedings
Johannes Grotendorst (Editor)
Winter School, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 1
ISBN 3-00-005618-1, February 2000, 562 pages
out of print

Modern Methods and Algorithms of Quantum Chemistry - Poster Presentations
Johannes Grotendorst (Editor)
Winter School, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 2
ISBN 3-00-005746-3, February 2000, 77 pages
out of print

Modern Methods and Algorithms of Quantum Chemistry - Proceedings, Second Edition
Johannes Grotendorst (Editor)
Winter School, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 3
ISBN 3-00-005834-6, December 2000, 638 pages
out of print

Nichtlineare Analyse raum-zeitlicher Aspekte der hirnelektrischen Aktivität von Epilepsiepatienten
Jochen Arnold
NIC Series Volume 4
ISBN 3-00-006221-1, September 2000, 120 pages

Elektron-Elektron-Wechselwirkung in Halbleitern: Von hochkorrelierten kohärenten Anfangszuständen zu inkohärentem Transport
Reinhold Lövenich
NIC Series Volume 5
ISBN 3-00-006329-3, August 2000, 146 pages

Erkennung von Nichtlinearitäten und wechselseitigen Abhängigkeiten in Zeitreihen
Andreas Schmitz
NIC Series Volume 6
ISBN 3-00-007871-1, May 2001, 142 pages

Multiparadigm Programming with Object-Oriented Languages - Proceedings
Kei Davis, Yannis Smaragdakis, Jörg Striegnitz (Editors)
Workshop MPOOL, 18 May 2001, Budapest
NIC Series Volume 7
ISBN 3-00-007968-8, June 2001, 160 pages

Europhysics Conference on Computational Physics - Book of Abstracts
Friedel Hossfeld, Kurt Binder (Editors)
Conference, 5 - 8 September 2001, Aachen
NIC Series Volume 8
ISBN 3-00-008236-0, September 2001, 500 pages

NIC Symposium 2001 - Proceedings
Horst Rollnik, Dietrich Wolf (Editors)
Symposium, 5 - 6 December 2001, Forschungszentrum Jülich
NIC Series Volume 9
ISBN 3-00-009055-X, May 2002, 514 pages

Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms - Lecture Notes
Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
NIC Series Volume 10
ISBN 3-00-009057-6, February 2002, 548 pages

Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms- Poster Presentations
Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
NIC Series Volume 11
ISBN 3-00-009058-4, February 2002, 194 pages

Strongly Disordered Quantum Spin Systems in Low Dimensions: Numerical Study of Spin Chains, Spin Ladders and Two-Dimensional Systems
Yu-cheng Lin
NIC Series Volume 12
ISBN 3-00-009056-8, May 2002, 146 pages

Multiparadigm Programming with Object-Oriented Languages - Proceedings
Jörg Striegnitz, Kei Davis, Yannis Smaragdakis (Editors)
Workshop MPOOL 2002, 11 June 2002, Malaga
NIC Series Volume 13
ISBN 3-00-009099-1, June 2002, 132 pages

Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms - Audio-Visual Lecture Notes
Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
NIC Series Volume 14
ISBN 3-00-010000-8, November 2002, DVD

Numerical Methods for Limit and Shakedown Analysis
Manfred Staat, Michael Heitzer (Eds.)
NIC Series Volume 15
ISBN 3-00-010001-6, February 2003, 306 pages

Design and Evaluation of a Bandwidth Broker that Provides Network Quality of Service for Grid Applications
Volker Sander
NIC Series Volume 16
ISBN 3-00-010002-4, February 2003, 208 pages

Automatic Performance Analysis on Parallel Computers with SMP Nodes
Felix Wolf
NIC Series Volume 17
ISBN 3-00-010003-2, February 2003, 168 pages

Haptisches Rendern zum Einpassen von hochaufgelösten Molekülstrukturdaten in niedrigaufgelöste Elektronenmikroskopie-Dichteverteilungen
Stefan Birmanns
NIC Series Volume 18
ISBN 3-00-010004-0, September 2003, 178 pages

Auswirkungen der Virtualisierung auf den IT-Betrieb
Wolfgang Gürich (Editor)
GI Conference, 4 - 5 November 2003, Forschungszentrum Jülich
NIC Series Volume 19
ISBN 3-00-009100-9, October 2003, 126 pages

NIC Symposium 2004
Dietrich Wolf, Gernot Münster, Manfred Kremer (Editors)
Symposium, 17 - 18 February 2004, Forschungszentrum Jülich
NIC Series Volume 20
ISBN 3-00-012372-5, February 2004, 482 pages

Measuring Synchronization in Model Systems and Electroencephalographic Time Series from Epilepsy Patients
Thomas Kreutz
NIC Series Volume 21
ISBN 3-00-012373-3, February 2004, 138 pages

Computational Soft Matter: From Synthetic Polymers to Proteins - Poster Abstracts
Norbert Attig, Kurt Binder, Helmut Grubmüller, Kurt Kremer (Editors)
Winter School, 29 February - 6 March 2004, Gustav-Stresemann-Institut Bonn
NIC Series Volume 22
ISBN 3-00-012374-1, February 2004, 120 pages

Computational Soft Matter: From Synthetic Polymers to Proteins - Lecture Notes
Norbert Attig, Kurt Binder, Helmut Grubmüller, Kurt Kremer (Editors)
Winter School, 29 February - 6 March 2004, Gustav-Stresemann-Institut Bonn
NIC Series Volume 23
ISBN 3-00-012641-4, February 2004, 440 pages

Synchronization and Interdependence Measures and their Applications to the Electroencephalogram of Epilepsy Patients and Clustering of Data
Alexander Kraskov
NIC Series Volume 24
ISBN 3-00-013619-3, May 2004, 106 pages

High Performance Computing in Chemistry
Johannes Grotendorst (Editor)
Report of the Joint Research Project:
High Performance Computing in Chemistry - HPC-Chem
NIC Series Volume 25
ISBN 3-00-013618-5, December 2004, 160 pages

Zerlegung von Signalen in unabhängige Komponenten: Ein informationstheoretischer Zugang
Harald Stögbauer
NIC Series Volume 26
ISBN 3-00-013620-7, April 2005, 110 pages

Multiparadigm Programming 2003
Joint Proceedings of the
3rd International Workshop on Multiparadigm Programming with Object-Oriented Languages (MPOOL'03)
and the
1st International Workshop on Declarative Programming in the Context of Object-Oriented Languages (PD-COOL'03)
Jörg Striegnitz, Kei Davis (Editors)
NIC Series Volume 27
ISBN 3-00-016005-1, July 2005, 300 pages

Integration von Programmiersprachen durch strukturelle Typanalyse und partielle Auswertung
Jörg Striegnitz
NIC Series Volume 28
ISBN 3-00-016006-X, May 2005, 306 pages

OpenMolGRID - Open Computing Grid for Molecular Science and Engineering
Final Report
Mathilde Romberg (Editor)
NIC Series Volume 29
ISBN 3-00-016007-8, July 2005, 86 pages

GALA Grünenthal Applied Life Science Analysis
Achim Kless and Johannes Grotendorst (Editors)
NIC Series Volume 30
ISBN 3-00-017349-8, November 2006, 204 pages

Computational Nanoscience: Do It Yourself!
Lecture Notes
Johannes Grotendorst, Stefan Blügel, Dominik Marx (Editors)
Winter School, 14. - 22 February 2006, Forschungszentrum Jülich
NIC Series Volume 31
ISBN 3-00-017350-1, February 2006, 528 pages

NIC Symposium 2006 - Proceedings
G. Münster, D. Wolf, M. Kremer (Editors)
Symposium, 1 - 2 March 2006, Forschungszentrum Jülich
NIC Series Volume 32
ISBN 3-00-017351-X, February 2006, 384 pages

Parallel Computing: Current & Future Issues of High-End Computing
Proceedings of the International Conference ParCo 2005
G.R. Joubert, W.E. Nagel, F.J. Peters,
O. Plata, P. Tirado, E. Zapata (Editors)
NIC Series Volume 33
ISBN 3-00-017352-8, October 2006, 930 pages

From Computational Biophysics to Systems Biology 2006 Proceedings
U.H.E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann (Editors)
NIC Series Volume 34
ISBN-10 3-9810843-0-6, ISBN-13 978-3-9810843-0-6,
September 2006, 224 pages

Dreistufig parallele Software zur Parameteroptimierung von Support-Vektor-Maschinen mit kostensensitiven Gütemaßen
Tatjana Eitrich
NIC Series Volume 35
ISBN 978-3-9810843-1-3, March 2007, 262 pages

From Computational Biophysics to Systems Biology (CBSB07) Proceedings
U.H.E. Hansmann, J. Meinke, S. Mohanty, O. Zimmermann (Editors)
NIC Series Volume 36
ISBN 978-3-9810843-2-0, August 2007, 330 pages

Parallel Computing: Architectures, Algorithms and Applications - Book of Abstracts
Book of Abstracts, ParCo 2007 Conference, 4. - 7. September 2007
G.R. Joubert, C. Bischof, F. Peters, T. Lippert, M. Bücker, P. Gibbon, B. Mohr (Eds.) NIC Series Volume 37
ISBN 978-3-9810843-3-7, August 2007, 216 pages

Parallel Computing: Architectures, Algorithms and Applications - Proceedings
Proceedings, ParCo 2007 Conference, 4. - 7. September 2007
C. Bischof, M. Bücker, P. Gibbon, G.R. Joubert, T. Lippert, B. Mohr, F. Peters (Eds.)
NIC Series Volume 38
ISBN 978-3-9810843-4-4, December 2007, 830 pages

All volumes are available online at

http:// www.fz-juelich.de/nic-series/.